



PHD

Establishing and Ensuring the Health, Fitness and Operational Performance of UK Fire & Rescue Service Personnel

Stevenson, Rich

Award date:
2018

Awarding institution:
University of Bath

[Link to publication](#)

Alternative formats

If you require this document in an alternative format, please contact:
openaccess@bath.ac.uk

Copyright of this thesis rests with the author. Access is subject to the above licence, if given. If no licence is specified above, original content in this thesis is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC-ND 4.0) Licence (<https://creativecommons.org/licenses/by-nc-nd/4.0/>). Any third-party copyright material present remains the property of its respective owner(s) and is licensed under its existing terms.

Take down policy

If you consider content within Bath's Research Portal to be in breach of UK law, please contact: openaccess@bath.ac.uk with the details. Your claim will be investigated and, where appropriate, the item will be removed from public view as soon as possible.

Establishing and Ensuring the Health, Fitness and Operational Performance of UK Fire & Rescue Service Personnel

Richard Daniel Mark Stevenson

A thesis submitted for the degree of Doctor of Philosophy

University of Bath

Department for Health

November 2017

COPYRIGHT

Attention is drawn to the fact that copyright of this thesis/portfolio rests with the author and copyright of any previously published materials included may rest with third parties. A copy of this thesis/portfolio has been supplied on condition that anyone who consults it understands that they must not copy it or use material from it except as permitted by law or with the consent of the author or other copyright owners, as applicable.

This thesis/portfolio may be made available for consultation within the University Library and may be photocopied or lent to other libraries for the purposes of consultation with effect from _____(date).

Signed on behalf of the Faculty/School of.....

Abstract

Firefighting is a strenuous occupation requiring high levels of physical fitness. Inadequate levels of fitness can put firefighters at risk of overexertion and injury. Therefore ensuring that firefighters maintain role specific fitness levels throughout their career is critical to both firefighter and public safety. The aim of this thesis was to investigate the minimum cardiorespiratory, strength and muscular endurance demands of UK firefighting and to recommend minimum physical employment standards to ensure the operational effectiveness and safety of firefighting personnel working in the UK fire & rescue service.

The first study developed a task analysis protocol to identify the minimum acceptable performance requirements of the critical and most physically demanding tasks in UK firefighting, identifying 2 distinct roles (firefighter and incident commander) with 8 critical tasks identified for firefighter and 2 for the incident commander role. The second study investigated the physical demands of performing these critical tasks to the minimum acceptable performance requirement. Cardiorespiratory fitness standards were derived for those undertaking both firefighting and incident command roles.

Following this, the validity and reliability of a firefighter simulation test was assessed to determine its appropriateness as a criterion test of operational fitness. Whilst there was a strong inverse correlation between the test completion time and cardiorespiratory fitness and the simulation was highly reliable, the error associated with the simulation suggests that it may not be suitable to use as a standalone fitness test and should be used in conjunction with gym-based cardiorespiratory fitness assessments. The final study assessed the sensitivity and specificity of common and replicable gym-based physical ability tests to predict firefighting performance in order to recommend strength and muscular endurance standards. Each of the gym-based physical ability tests and associated standards were effective at predicting effective firefighting performance.

Published Journal Articles

Stevenson, R. D., Siddall, A. G., Turner, P. F., & Bilzon, J. L. (2017). Physical Employment Standards for UK Firefighters: Minimum Muscular Strength and Endurance Requirements. *Journal of Occupational and Environmental Medicine*, 59(1), 74-79.

Stevenson, R. D., Siddall, A. G., Turner, P. F., & Bilzon, J. L. (2016). A Task Analysis Methodology for the Development of Minimum Physical Employment Standards. *Journal of Occupational and Environmental Medicine*, 58(8), 846-851.

Siddall, A. G. and Stevenson, R. D., Turner, P. F., Stokes, K. A., & Bilzon, J. L. (2016). Development of role-related minimum cardiorespiratory fitness standards for firefighters and commanders. *Ergonomics*, 59(10), 1335-1343.

Published Abstracts

Stevenson, R. D., Siddall, A. G., Turner, P. F., Stokes, K. A., & Bilzon, J. L. (2014). The Metabolic and Cardiovascular Demands of UK Firefighting Tasks: 3471. Board #115. *Medicine & Science in Sports & Exercise*, 46:930.

Stevenson, R. D., Siddall, A. G., Turner, P. F., Stokes, K. A., & Bilzon, J. L. (2016). Task Analysis for the Development of Minimum Physical Employment Standards for Physically Demanding Occupations. *Applied Physiology, Nutrition, and Metabolism*, 41(6 (Suppl. 2)), S194.

Conference Presentations

Stevenson, R. D., Siddall, A. G., Turner, P. F., & Bilzon, J. L. (2015). A Task Analysis Methodology for the Development of Minimum Physically Demanding Occupations. *Second International Conference on Physical Employment Standards*, Canmore, Alberta, Canada.

Stevenson, R. D., Siddall, A. G., Turner, P. F., & Bilzon, J. L. (2014). The Metabolic and Cardiovascular Demands of UK Firefighting. *American College of Sports Medicine World Congress*, Orlando, Florida, USA.

Acknowledgements

I would like to express my gratitude to everyone who has supported me throughout the course of this project. Firstly to Professor Kevin Sykes for planting the seed and giving me the confidence to embark on this journey. To Professor James Bilzon, thank for your friendship and supervision throughout this project and for being a constant source of stability through the scientific and particularly the political turbulence as well as for the opportunities to present my work to international audiences. To Dr Andy Siddall, thank you for your continuous guidance, patience and support over the many rounds of flat white coffees. To Professor Keith Stokes and Dr Tom Nightingale for the direction and support when it was needed.

I would also like to thank the UK fire and rescue service and in particular South Wales fire and rescue service and the Chief Fire Officers Association for their unwavering support but mostly the determination to develop fitness standards under incredibly challenging political times. Special thanks must go to Mark Malson, Justin Johnston, Phil Turner and Des Pritchard, everyone on the FireFit steering group and of course to all of the participants and facilitators across the UK who made the project achievable. Without you all this would not have happened. A special thank you must also go to Gareth Green, my trusted friend and colleague for his incredible support who has been on this journey with me longer than anyone else. You truly deserve some kind of medal.

I would also like to thank my friends and family. To Mum and Dad, your support is endless and I cannot thank you enough for the sacrifices you have made for me. To my wife, Anna and son Ben, to me you are both truly amazing and a constant source of inspiration. Anna, thank you for your never-ending support and encouragement. To my brothers Al and Grae and close friends Jake, Cathy, Dom and others for having put up with the endless venting of issues relating to firefighter fitness. Thank you all.

Statement of Authorship

This research project included a number of contributors and are detailed below:

Conceptualisation	This concept of this project was initiated by R.D.M. Stevenson
Realisation	R.D.M. Stevenson, A. Siddall, P. Turner & J. Bilzon
Data Collection	R.D.M. Stevenson, A. Siddall, P. Turner
Chapter 1	Written by R.D.M. Stevenson
Chapter 2	Written by R.D.M. Stevenson and reviewed by J. Bilzon
Chapter 3	Written by R.D.M. Stevenson. The scientific paper (Stevenson et al. 2016) was written by R.D.M. Stevenson (1 st author) along with A. Siddall, P. Turner, J. Bilzon.
Chapter 4	Written by R.D.M. Stevenson. The scientific paper (Siddall and Stevenson et al., 2016) was co-written by R.D.M. Stevenson and A. Siddall (1 st authors) along with P. Turner, J. Bilzon & K. Stokes.
Chapter 5	Written by R.D.M. Stevenson along with A. Siddall, P. Turner, J. Bilzon.
Chapter 6	Written by R.D.M. Stevenson. The scientific paper (Stevenson et al., 2017) was written by R.D.M. Stevenson (1 st author) along with A. Siddall, P. Turner & J. Bilzon
Chapter 7	Written by R.D.M. Stevenson and reviewed by J. Bilzon

Table of Contents

Abstract	2
Journal Publications	3
Conference Presentations	4
Acknowledgements	5
Statement of Authorship	6
Table of Contents	7
List of Figures	10
List of Tables	11
Abbreviations	13
CHAPTER 1	14
Introduction	15
CHAPTER 2	17
Review of Literature	18
2.1 The importance of physical fitness for demanding occupations.....	18
2.1.1 A historical perspective	18
2.1.2 Firefighter fatalities	20
2.2 The physical demands of firefighting.....	22
2.2.1 The role of the UK Fire & Rescue Service	22
2.2.2 The critical and most physically demanding tasks	23
2.2.3 Cardiorespiratory demands	27
2.2.4 Muscular strength and endurance requirements.....	33
2.3 Establishing physical employment standards	37
2.3.1 Legal considerations and physical employment standards	38
2.3.2 Procedures for establishing physical employment standards.....	42
2.3.3 Physical employment tests	48
2.3.4 Setting standards.....	51
CHAPTER 3	56
A Task Analysis Methodology for the Development of Physical Employment	
Standards	57
3.1 Introduction.....	57
3.2 Methods	59

3.3 Results	64
3.4 Discussion.....	67
CHAPTER 4	72
Development of Role-related Minimum Cardiorespiratory Fitness Standards for Firefighters and Commanders.....	73
4.1 Introduction.....	73
4.2 Methods	74
4.3 Results	81
4.4 Discussion.....	88
CHAPTER 5	94
Establishing the Validity and Reliability of a Firefighting Simulation Test.....	95
5.1 Introduction.....	95
5.2 Methods	96
5.3 Results	100
5.4 Discussion.....	102
CHAPTER 6	106
Development of Minimum Muscular Strength and Endurance Requirements	107
6.1 Introduction.....	107
6.2 Methods	108
6.3 Results	113
6.4 Discussion.....	117
CHAPTER 7	121
Discussion.....	122
7.1 General Discussion.....	122
7.1.1 What are the critical and most physically demanding generic tasks in the UK fire and rescue service?	122
7.1.2 What are the role-related minimum cardiorespiratory fitness standards for firefighters and commanders?	124
7.1.3 What is the validity and reliability of a firefighting simulation test (FFST)?	125

7.14 What are the minimum muscular strength and endurance requirements for firefighters?	126
7.15 Practical implications.....	128
7.16 Supporting research	128
7.2 Future Directions.....	129
7.2.1 Specialist roles.....	129
7.2.2 Trainability of UK Firefighters	130
7.2.3 The effects of physical fitness on decision making and mental ability in firefighters.....	131
References	132

List of Figures

Figure 2.1 Decision theory matrix (contingency table)	53
Figure 3.1 Task analysis consultation process using convened meetings with technical (subject-matter experts) and stakeholder panels.....	61
Figure 5.1 Relationship between completion time (seconds) on the firefighting simulation versus VO_2max ($n = 69$, $r = -0.734$, $r^2 = 0.539$), with line of best fit.	101
Figure 5.2 Test-retest reliability of the firefighter simulation test ($n = 22$, $r = -0.842$, $r^2 = 0.708$).	102
Figure 6.1 Individual performances (Pass/Fail) in the criterion ladder lift task against 1RM in the seated shoulder press test (A), and corresponding ROC curve (B) derived from these data.....	115
Figure 6.2 Individual performances (Pass/Fail) in the criterion ladder lower task against 1RM in seated maximal rope pull-down (A) and corresponding ROC curve (B) derived from these data.....	116
Figure 6.3 Individual performances (Pass/Fail) in criterion ladder extension task against the number of repetitions to failure during the repeated rope pull-down at $35 \text{ pulls} \cdot \text{min}^{-1}$ (A) and corresponding ROC curve (B) derived from these data.....	116
Figure 7.1 Fitness management process for the UK fire and rescue services	128

List of Tables

Table 2.1 Summary of research investigating the cardio-respiratory demands of simulated firefighting tasks	30
Table 2.2 Correlation matrix of selected firefighting performance and common generic ability tests.....	35
Table 3.1 Speeds of each recorded video, for each task and corresponding voting options.....	66
Table 3.2 Technical panel choices, mean, mode, range and consensus scores with corresponding minimum acceptable work rates.....	66
Table 3.3 Summary of the practical steps undertaken.	67
Table 4.1 Participant Characteristics. Data are mean (\pm SD).	76
Table 4.2 Participant responses to the instruction received, authenticity and pace of each task as minimum acceptable performance.	81
Table 4.3 Metabolic demand, cardiovascular strain and perceived exertion for peak steady state during firefighting tasks for participants who completed all tasks successfully (n=47).....	84
Table 4.4 Peak steady state metabolic demand data from those completing each task successfully, organised by sex.....	85
Table 4.5 Metabolic demand, cardiovascular strain for peak steady state during simulated ‘incident commander’ duty within firefighting tasks.....	86
Table 5.1 Description of the consecutive components of the firefighting simulation .	98

Table 5.2 Study participant physical and performance characteristics (mean \pm SD). 100

Table 5.3 Reliability study participants' physical and performance characteristics (mean \pm SD)..... 101

Table 6.1 Performance on the criterion tasks and the corresponding maximal performances (mean \pm SD) during the physical ability test(s)..... 114

Abbreviations

1RM	One Repetition Max
ANOVA	Analysis Of Variance
BMI	Body Mass Index
CFOA	Chief Fire Officers Association
CHD	Coronary Heart Disease
CPAT	Canadian Physical Ability recruitment Test
CV	Coefficient of Variation
FFST	Firefighter Selection Test
GPT	Generic Predictive Test
HRR	Health Rate Reserve
KG	Kilogram
PES	Physical Employment Standard
PET	Physical Employment Test
PC	Police Constable
ROC	Receiver-Operating Characteristic
RPE	Ratings of Perceived Exertion
SD	Standard Deviation
SEE	Standard Error of Estimate
SP	Stakeholder Panel
TP	Technical Panel
TPT	Task-related Predictive Test
TST	Task Simulation Test
UK	United Kingdom
USA	United States of America
$\dot{V}O_2\text{max}$	Maximal Oxygen Uptake

CHAPTER 1

General Introduction

Introduction

The history of firefighting is reported to have begun in ancient Rome in the 3rd Century (The Firefighter Foundation). However it was not until 1824 that the first community Fire Service was established in Edinburgh and named the Edinburgh Fire Engine Establishment. Since then the role of firefighters and the fire service has evolved dramatically, yet despite continued advancements in equipment technology, building design and fire risk management strategies, firefighting remains a physically demanding occupation (Bos, Mol, Visser, & Frings-Dresen, 2004; Elsner, 2008; Holmer & Gavhed, 2007; Jamnik, Thomas, Shaw, & Gledhill, 2010; Von Heimburg, Rasmussen, & Medbo, 2006). As well as having to use heavy firefighting equipment, firefighters are required to respond to emergency situations within minutes, work in dangerous environments often involving heat, humidity and toxic smoke whilst wear heavy protective clothing. It is this unique interaction of time critical situations, the need to mobilise heavy equipment coupled with extreme environmental temperatures and cumbersome personal protective equipment that combine to exert substantial physical strain on firefighters (Office of the Deputy Prime Minister, 2004b). For these reasons, physical fitness is considered critical to the role of a firefighter.

The distinctive physiological demands coupled with the importance of firefighting to protect property and lives may be why research into the physicality of firefighting has been undertaken for almost half a century. This research has strengthened the understanding of the importance of physical fitness for firefighters and others working in physically demanding occupations. Whilst much of the earlier research linked improved physical fitness with the ability to carry out the work in a more effective and efficient manner (Davis, Dotson, & Maria, 1982; Kilbom, 1980; Lemon & Hermiston, 1977), more recent studies have identified the link between a lack of physical fitness and health risks to the employees when undertaking such demanding work (Baur, Christophi, & Kales, 2012; Kales, Soteriades, Christophi, & Christiani, 2007; Smith, Barr, & Kales, 2013; Yang et al., 2013).

The combined effect of this research has highlighted the importance of physical fitness and issues relating to substandard fitness capacities to both firefighters and members of the public. It is these links that have helped to shape how physical fitness contributes to workplace health and safety in physically demanding jobs. In many developed countries around the world, ensuring firefighters maintain appropriate levels of physical fitness is now seen as a part of an employer's duty of care to certify that firefighters are able to carry out their work both safely and effectively, thus ensuring firefighters safety and the safety of the public (Great Britain Parliament, 1974; Health & Safety Executive, 2010; Jamnik, Gumienak, & Gledhill, 2013).

However, despite a number of reports highlighting the importance of physical fitness in UK firefighters (Office of the Deputy Prime Minister, 2004b) and the need for regular fitness assessments (Stevenson, Wilsher, & Sykes, 2009), the physical demands of firefighting in the UK has not been quantified. This lack of scientific evidence has delayed the development of defensible national fitness standards and valid health & wellbeing programmes to support firefighters maintaining role specific fitness levels. Therefore the research in this thesis will aim to answer the following research questions:

- What are the critical and most physically demanding generic tasks in the UK fire and rescue service?
- What are the role-related minimum cardiorespiratory fitness standards for firefighters and commanders?
- What is the validity and reliability of a firefighting simulation test?
- What are the minimum muscular strength and endurance requirements for firefighters?

CHAPTER 2

Review of Literature

Review of Literature

2.1 The importance of physical fitness for demanding occupations

2.1.1 A historical perspective

The importance of physical conditioning to produce battle-ready warriors is reported to predate the Greek civilisation of the 1st Century B.C (East, 2013). However, it is the Greeks and in particular the Spartans that are credited for developing and formalising physical training regimes for their warriors (East, 2013). From the age of seven, Spartan boys were taken from their families and sent to live in military training camps where they were taught discipline, fighting, hunting and survival skills, as well as the endurance of pain (Cartledge, 2004). At the age of 20, these young adult Spartans completed their training and became soldiers of the state (Cartledge, 2004).

During the 1st and 2nd Centuries A.D. the Romans continued these structured training practices adopted by the Greeks, which required soldier training to bring recruits to the highest point of physical robustness by daily exercise (Stout, 1921). Initially, Roman soldiers were enrolled for one fighting campaign only, but the enlistment of professional soldiers soon followed with the introduction of primitive physical recruitment standards (Stout, 1921). Roman soldiers were required to be at least 5ft. 10 inches tall (equivalent to a modern day 5ft. 8 inches), ‘have observant eyes, a broad chest, muscular shoulders, strong arms, not too extended a waist measure or with superfluous flesh but hard and knotted with muscles’ (Stout, 1921). Whilst this was a noteworthy step towards physical recruitment criteria, these standards were subjective, difficult to standardise and would have been challenging to implement in a consistent manner.

Through the Middle Ages, the concepts of physical training for military preparedness remained relatively unchanged with fitness focusing on maintaining combat skills and improving dynamism. In the 18th Century, Dr George Turnbull, a Lieutenant Colonel in the American Revolutionary War was credited with broadening the understanding of physical fitness and identified the need for physical activity to ‘invigorate the soul as well as the body to produce courage, firmness and vigour’ (East,

2013). He was also reported to be one of the first to have written on the importance of physical exercise for successful military service. However, perhaps the single most influential step in the structured physical training of soldiers began in 1793 with the publication of *Gymnastik Fuer Die Jugend* (Gymnastics for Youth) by a German school teacher (Johann Christoph Friedrich GutsMuths), who was credited with revolutionising the understanding of physical conditioning by developing a core curriculum of novel callisthenic exercises and workouts (East, 2013).

By the 19th century, physical fitness and particularly the concepts of gymnastic exercise were well recognised as being critical for effective military training across Europe and North America and many of the worlds armies began to adopt gymnastic style training as a way of increasing the fitness of their soldiers (East, 2013; Knapik, 1989; Knapik & East, 2014). In 1858, a Scottish gymnastics teacher Archibald McLaren developed a system of military gymnastics for the British Army and went on to become the Director of Army Gymnastics Staff. In the same year, the U.S. Army implemented the first official physical selection tests within the United States Military Academy (East, 2013). From the middle of the 19th century onwards, physical selection criteria and structured physical training regimes began to form an integral part of the selection and training of military personnel across the world. Whilst there have been many changes in military technology and combat tactics, physical fitness is still recognised as being critical to the success of military operations.

Other physically demanding public service jobs can also be dated back as far as Roman times with the history of firefighters being reported to have begun in ancient Rome in the 3rd century (The Firefighter Foundation, 2017). The first collection of 500 firefighters was reportedly led by a Roman general by the name of Marcus Licinius Crassus. Firefighters patrolled the streets of Rome, and when needed lined up with buckets from the nearest water source to help control the fire. The first organised firefighters in Britain were understood to have been adopted following the Roman invasion of AD43, though practices were reported to be rudimentary and often unreliable (The Firefighter Foundation, 2017). However, the great fire of London in 1666 changed this and led the way for a more structured fire service in the UK.

In 1667, an insurance company called the Fire Office was established in London employing small teams of firefighters. Soon after, other companies began to offer similar services across the UK, but it was not until 1824 when the first community Fire Service in the world was established in Scotland and named the Edinburgh Fire Engine Establishment (The Firefighter Foundation, 2017). In 1938, just before the Second World War, the UK's National Fire Service was created to help ensure uniformity in the way in which fire cover was provided. However, the importance of physical fitness for firefighters in the UK was not officially documented until 1978 when the Fire Service Appointments and Promotion Regulations indicated that 'pre-employment medical assessments should include an aerobic capacity step test to a level that would enable the person to undertake firefighting duties' (Haisman, 1996). Much like the military, the importance of physical fitness in the fire service initially focused on developing fit personnel to be able to do the job rather than ensuring the safety of employees.

2.1.2 Firefighter fatalities

However, in 2007 a key scientific study published in the U.S. changed the thinking on physical fitness and added a whole new dimension to its importance for firefighters and other workers in physically demanding occupations. This research, which analysed firefighter fatalities across the whole of the U.S. between 1994 and 2004, found that the most considerable risk to U.S. firefighters was from coronary heart disease (CHD) (Kales et al., 2007). It reported that over the 10 year period, the leading cause of on-duty firefighter death (with 45% of all on-duty fatalities), was from heart attack caused by CHD (Kales et al., 2007). Thirty-two per cent of these fatalities occurred whilst undertaking fire suppression activities, however, returning from an alarm (17.4%), responding to an alarm (13.4%) and physical training (12.5%) were the next most likely activities to be associated with death from this cause (Kales et al., 2007), all of which involved acute cardiovascular arousal and / or strain.

The authors concluded that an increased cardiovascular stimulation associated with alarm response and the cardiovascular demands associated with fire suppression and physical training triggering the cardiovascular event was the most likely explanation for these increases in the number of fatalities. The authors also went on to postulate that

this may be, in part, because ‘many firefighters may have underlying cardiovascular risk factors and lack adequate fitness’ (Kales et al., 2007). Cardiovascular risk can be elevated in firefighters because of factors such as shift work, exposure to toxic gases, extreme physical and emotional demands, obesity and low physical fitness (Baur, Christophi, & Kales, 2012; Plat, Frings-Dresen, & Sluiter, 2011; Roberts, 2002; Smith et al., 2013; Soteriades, Smith, Tsismenakis, Baur, & Kales, 2011; Willich et al., 1993; Yang et al., 2013). One study identified that 115 firefighter recruits from one US fire department had a mean cardiorespiratory fitness level of 35 ml·kg·min⁻¹ at the start their careers, which is approximately 20% below that required for safe and effective firefighting (Roberts, O’Dea, Boyce, & Mannix, 2002).

In the UK, evidence also suggests that UK firefighters may be subjected to similar cardiovascular risks. In a review of firefighter fatalities spanning 30 years between 1978 and 2008, 36 deaths (30% of the total) were “attributed to natural causes, which were later described as being ‘generally heart attacks’, taking place either at operational incidents or shortly afterwards” (Labour Research Department, 2008). As the physical demands of undertaking UK firefighting to the minimum acceptable standard has not been quantified, firefighters are not provided with any official physical training guidance to maintain minimum fitness standards for their role(s). Also whilst the training and assessment of many fire and rescue based competencies have been standardised across the UK, the structuring of fitness assessments has not. A survey of fitness testing practices taken at a national fitness conference in 2009 by the FireFit Steering Group identified that there was a wide variation in the way Fire & Rescue Services were delivering routine fitness assessments (FireFit Steering Group, 2009). Only 14% of Services administered fitness tests every 6 months in line with Her Majesty’s Chief Inspector of Fire Services recommendations (Her Majesty’s Fire Service Inspectorate, 2000), whilst thirty-seven percent also indicated that they applied potentially discriminatory age and/or gender modified standards with 20% of Services were still not performing any ongoing fitness assessments at all.

2.2 The physical demands of firefighting

2.2.1 The role of the UK Fire & Rescue Service

Until relatively recently, the UK fire services had delivered a predominantly reactive service in the event of an emergency to protect the public from the dangers of fire. However, in more recent years, the role of the fire service has changed greatly with firefighters having to respond to a much wider range of emergencies. The first significant change to the role of the fire service started in the 1980's where firefighters began to respond to road traffic collisions (Office of the Deputy Prime Minister, 2003). From the 1990's many of the UK's fire services changed their names to fire and 'rescue' services to reflect their diversified role (Office of the Deputy Prime Minister, 2003). However, despite changes to the types of emergencies firefighters attended, it was not until the turn of the 21st century that fire and rescue services moved from the traditional reactive position of fighting fires and responding to emergencies, to a much more proactive culture of preventing them (Office of the Deputy Prime Minister, 2003). In 2004 the Fire & Rescue Services Act set out new legislation with specific duties for fire and rescue services to promote fire safety and to respond to the needs of the community in respect to the specific risks that they may face (Office of the Deputy Prime Minister, 2003). This concept became widely known as community fire safety and it is these proactive initiatives that have been primarily credited for the reduction in fires in the UK over the last decade (Home Office, 2016).

Since the turn of the century, the UK fire and rescue services has also had to respond to a number of government reports identifying the need for change in the way they delivered its service (Office of the Deputy Prime Minister, 2000). As well as attending the commonplace emergencies such as fires and road traffic accidents, fire and rescue services across the UK have extended their remit further to respond to additional emergencies, this time including natural disasters involving severe weather conditions, major transport incidents, chemical, biological and nuclear threats, collapsed structures as well as the growing threat from terrorism (Office of the Deputy Prime Minister, 2003). In 2001, firefighters also began attending medical emergencies, which have traditionally been reserved for the ambulance services. Although slow in its uptake across the UK, firefighters continue to work more closely year on year with ambulance service staff to provide specific emergency medical response. In 2009, official figures

indicated that UK firefighters responded to 11,000 emergency medical calls, whereas by 2013 this number had increased by 36% to 15,000 (Home Office, 2016).

With the ever widening remit of the fire and rescue service in the UK, coupled with the reduction in actual fire calls, firefighters are less and less likely to attend a serious fire incident. Whilst preventing these serious fires is highly desirable, firefighters are less frequently exposed to the extreme physical demands of these incidents and a government report published in 2004 indicated that the frequency of exposures to the job appeared to be insufficient to develop and maintain role specific fitness levels (Office of the Deputy Prime Minister, 2004b). As the exposure to these physically demanding tasks reduces, the risk of injury from over-exertion can perversely increase if role related fitness levels are not maintained.

2.2.2 The critical and most physically demanding tasks

In contrast to the shifting role of the fire and rescue service, many of the traditional firefighting tasks have remained relatively unchanged over the years. Whilst there have been technological advancements in protective clothing, equipment design, and risk management strategies, firefighters still apply offensive tactics to protect property and save lives from the dangers of fire and other risks. In practical terms, this often involves carrying heavy firefighting equipment to the scene of a fire and rescuing casualties whilst wearing cumbersome personal protective equipment and breathing apparatus. Whilst understanding all aspects of the role of a firefighter are of interest, it is arguably the critical and most physically demanding tasks that are of greatest importance to ensure that these workers can perform these components of their work safely and effectively, thus fulfilling their public safety role. The combination of these high physical demands, the extreme occupational environments and the interaction of insulating personal protective equipment has intrigued occupational physiologists and ergonomists, leading to a significant body of investigations into the demands of critical firefighting tasks, over a number of decades.

In the Netherlands one study attempted to identify the critical and most demanding tasks undertaken by Dutch firefighters by describing the quantity, duration

and frequency of firefighting incidents through direct observation in fire stations scattered across the country (Bos et al., 2004). The study, which included 222 volunteer and professional firefighters from 20 fire stations across the Netherlands analysed the work tasks completed by employees during eighty-three 24-hour firefighting shifts. The frequency and duration of work incidents, including the specific physical actions completed during these incidents were recorded. The results indicated that on average each shift had only 1.5 incidents (range 0-7), with a mean turn out time of 88 minutes (range 11-481 minutes) per call out. The most common activities recorded, in order of frequency were lifting / carrying equipment, stooping, kneeling and squatting, pulling and dragging equipment, walking up and down stairs and running and crawling. However, activities including swimming, diving and hoisting equipment did not occur during the 24-hour shifts, despite them being recognised activities involved in firefighting. Identifying these unrecorded duties in this manner is an example of where direct observation may be limited particularly in jobs where activities are performed infrequently or haphazardly (Constable & Palmer, 2000). In order to fully understand the typical duration of an infrequent incident in this way, considerable observational time of weeks, months or even years may be needed, which would be unrealistic to perform. Furthermore, observing firefighters at work during operational scenarios may pose a significant danger to the observer.

In Canada, critical firefighting tasks are described in the Canadian Physical Ability recruitment Test (CPAT) developed by the International Association of Fire Chiefs. The test is described as a *bone fida* occupational requirement for firefighter applicants wishing to become professional firefighters in Canada. Whilst the original task analysis research is not reported within the peer-reviewed literature, the critical firefighting tasks are described in detail in other work (Williams-Bell, Villar, Sharratt, & Hughson, 2009). The critical tasks established for Canadian firefighting include stair climbing with equipment, dragging hose, carrying equipment, raising and extending fire service ladders, forcible entry into a building followed by searching for and rescuing casualties. Other Canadian research establishing pre-employment standards for the Canadian Forces and Department of Defence firefighters has identified similar tasks (Rogers, Docherty, & Petersen, 2014).

In 2005, the state of Florida Fire Marshall commissioned a task analysis of career and volunteer firefighters to determine the essential job tasks, skills and abilities necessary for successful firefighting performance in U.S firefighters (Industrial Organisational Solutions, 2015). The comprehensive study, which involved over 150 fire departments in the state of Florida, used the results of the analysis to develop minimum operational training requirements to ensure the technical competence of their operational staff. The analysis used job interviews, observations and questionnaires and ranked activities in terms of their importance for the role. The analysis revealed that 18 discrete physical abilities were identified and considered highly relevant to the role of firefighter by the subject matter experts questioned in the study. A range of physical movements including crawling, ducking under and manoeuvring around obstacles were considered important and tasks including wearing personal protective equipment and breathing apparatus, dragging and carrying hose and other firefighting equipment, searching for and rescuing victims and other firefighters were identified as the most physically demanding aspect of the role (Industrial Organisational Solutions, 2015).

In Australia, a large regional fire and rescue service undertook a comprehensive analysis of the critical firefighting tasks of urban firefighters in 2015 for the development of firefighter selection tests (Taylor, Fullagar, Mott, Sampson, & Groeller, 2015). The researchers utilised many of the most current best practice methods and used interviews and focus group meetings with 106 firefighters to identify 31 physically demanding tasks before developing and circulating a workforce survey to rate these firefighting activities in terms of importance, frequency, duration and perceived difficulty. Although only 25% of the total workforce responded to the survey, this is similar to other large workforce task analysis surveys (Office of the Deputy Prime Minister, 2004a) and a total of 989 responses were received. This also highlights one of the difficulties associated with using survey data in an occupational setting. The researchers used a decision-analysis approach to evaluate and filter the survey responses thus eliminating certain tasks that were duplicated by other activities. Through a series of cross-checking procedures, the number of critical tasks was reduced to 15 distinct tasks which included carrying equipment, dragging, rolling out and using 38 and 70 mm hose, climbing stairs, rescuing casualties, using ladders and other firefighter specific equipment (Taylor et al., 2015).

In the UK, large-scale projects focused on the physical elements of firefighting were initiated in 2002 to develop physical selection tests (Blacker et al., 2015) and 2004 to provide occupational advice for medical advisers on the recruitment of fire and rescue service employees, respectively (Office of the Deputy Prime Minister, 2004a). As part of this work, a detailed task analysis was undertaken with 18 distinct fire service roles identified and separated into four clustered groups which were referred to as Operational, Management / Safety, Control and Training positions. Operational firefighting roles were identified as being the most physically active with firefighters performing more than 20 distinct postures and movements whilst at work. These included activities such as running, climbing, squatting and stooping and the majority of operational staff reported engaging in these activities ‘many times per day’ (Office of the Deputy Prime Minister, 2004a).

The analysis identified that the most hands on operational roles (Firefighter, Crew Manager and Watch Manager) required a wide range of physical competencies to perform even the most commonplace activities. Other, more demanding activities, which occurred less frequently, were nonetheless still considered critical to these positions. It was identified that carrying firefighting hose, lifting and extending ladders and using cutting equipment weighing up to 40kg were important and some of the most physically demanding activities performed. Additionally, rescuing casualties whilst wearing protective equipment were considered a critical part of the firefighter role. In the development of physical employment standards (PES), Blacker et al. (2015) used the results of this task analysis and using subject matter experts in focus groups undertook a comprehensive analysis of the peer reviewed literature and current working practices of firefighting personnel and identified 7 distinct firefighting tasks critical to the generic firefighting role. Blacker et al. (2015) also developed both criterion simulations encompassing these physical demands as well as selection tests and standards that predicted performance on the criterion test. The tests simulations involved the following firefighting activities:

- Working at rural fires
- Working at domestic fires

- Lifting fire service ladders
- Extending fire service ladders
- Climbing fire service ladders
- Assembling firefighting equipment
- Working in enclosed spaces

In conclusion, it is clear that there are a number of similarities in many countries around the world in relation to the critical and most physically demanding activities that firefighters perform. Yet, despite various technological improvements in recent years, many of the critical tasks have remained the same. Having to move heavy equipment such as water pumps to provide a water supply and use fire service ladders to rescue casualties continues to form a fundamental part of the firefighting role across the world.

2.2.3 Cardiorespiratory demands

For a number of decades it has been recognised that cardio-respiratory fitness is an important factor for successful firefighting duties (Davis et al., 1982; Kilbom, 1980; Lemon & Hermiston, 1977). This seems intuitive given that exposure to operational firefighting scenarios requires sustained physical work, often for 20-30 min, which is often dictated by the capacity of self-contained breathing apparatus. Since this time, researchers have sought to investigate these physical demands and in a number of cases recommend minimum standards to guarantee successful firefighting performance and to help ensure firefighter and public safety (Bilzon, Scarpello, Smith, Ravenhill, & Rayson, 2001; Gledhill & Jamnik, 1992a; Lemon & Hermiston, 1977; Sothmann et al., 1990). However, despite significant technological advancements, the live fire environment still poses significant challenges to investigating these physical demands. Alternatively, researchers have turned to the next best approach by investigating firefighters performing simulated firefighting tasks.

Measuring the physical demand of physical activity or physical work is analysed by indirect calorimetry, a technique that measures inspired and expired gas flows including volumes and O₂ and CO₂ concentrations which have reported to have begun

almost 200 years ago (American College of Sports Medicine, 2010; McClave & Snider, 1992). For many years, expired air samples were collected from exercising individuals using collection or Douglas bags and the contents of the bag analysed at a short time after (Shephard, 1955). However, through technological advancement, the introduction of the computerised static systems and successively portable analyser allowed for the real time 'breath by breath' analysis in a range of sport and occupational settings (Howe, Matzko, Piasek, Pitsiladis, & Easton, 2014). Whilst some researchers have suggested that collection bags remain the gold standard for the measurement of expired air samples the portable metabolic analyser has been central to the understanding of metabolic demand of numerous activities and jobs including firefighting (Howe et al., 2014; Pinnington, Wong, Tay, Green, & Dawson, 2001).

However, it has been identified that all gas analysis systems, including portable metabolic analysers such as the Cosmed K4b² (Rome, Italy) suffer from sources of technical and / or biological error which can impact the accuracy of its measurements, therefore it is important to be able to measure the accuracy of such systems (Howe et al., 2014). Whilst some research investigating the accuracy of the K4^b has been equivocal (Duffield, Dawson, Pinnington, & Wong, 2004), possibly due to the large variety in exercise modes and intensities studied, other research have reported this system to be a reliable and valid mode for measuring oxygen consumption in a number of activities including walking, cycling, running (Duffield et al., 2004; Schrack, Simonsick, & Ferrucci, 2010). However, some have suggested that the accuracy of the K4^b can overestimate VO₂ measurements, particularly when measuring for longer durations lasting longer than 1 hour (Howe et al., 2014).

Whilst researchers have described the importance of cardio-respiratory fitness in both absolute values (Rayson, Holliman, & Belyavin, 2000) and also relative to body mass (Bilzon et al., 2001), much of the recommendations in relation to standards have been reported in relative maximum oxygen uptake (VO₂ ml·kg⁻¹·min⁻¹), partly because operational firefighters must support their own body mass while conducting physical work. The first known study to investigate the energy cost of completing simulated firefighting tasks was published in 1977. The researchers were able to overcome many

of the substantial practical issues with this type of work of measuring the oxygen demand of firefighting tasks though the use of expired air collection bags (Lemon & Hermiston, 1977). Subsequent research was undertaken using portable metabolic analysers described previously. Table 2.1 summarises key research projects investigating the cardio-respiratory demands of simulated firefighting tasks through the direct measurement of expired respiratory gases. The findings presented show a relatively large spectrum of mean cardio-respiratory demands of firefighting tasks ranging from $23 \text{ ml}\cdot\text{kg}\cdot\text{min}^{-1}$ (Bilzon et al., 2001) to $44 \text{ ml}\cdot\text{kg}\cdot\text{min}^{-1}$ (Holmer & Gavhed, 2007) with the subsequent recommended minimum standards ranging between $33.5 \text{ ml}\cdot\text{kg}\cdot\text{min}^{-1}$ (Sothmann et al., 1990) and $4.0 \text{ l}\cdot\text{min}^{-1}$ (equivalent to $50 \text{ ml}\cdot\text{kg}\cdot\text{min}^{-1}$ for an 80kg firefighter) (Von Heimburg et al., 2006). However factors such as the task design and duration, task type and the kind of pacing strategy used may explain some of the variability in some of these results.

Table 2.1. Summary of research investigating the cardio-respiratory demands of simulated firefighting tasks

Author	Year	Country	Sample	Analysis method	Pacing	Mean Task Duration	Oxygen demand	Minimum recommendation
Lemon & Hermiston	1977	Canada	20M	Collection bags	Paced	48 sec	9.2 - 10.2 METs	40 ml·kg ⁻¹ ·min ⁻¹
Sothmann et al.	1990	USA	Sub sample of 20M	Direct gas analysis	Normal firefighting pace	9 mins	Mean 30.5 ml·kg ⁻¹ ·min ⁻¹	33.5 ml·kg ⁻¹ ·min ⁻¹
Gledhill & Jamnik	1992	Canada	60	Direct gas analysis	Normal firefighting manner	25 sec – 3 mins	23 – 44 ml·kg ⁻¹ ·min ⁻¹	45 ml·kg ⁻¹ ·min ⁻¹
Bilzon et al.	2001	UK	34M / 15F	Direct gas analysis	Paced	4 mins	Peak 23 – 43 ml·kg ⁻¹ ·min ⁻¹	41 ml·kg ⁻¹ ·min ⁻¹
Holmer & Gavhed	2006	Sweden	15M	Direct gas analysis	As quickly as possible	22 mins	Mean 33.9 ml·kg ⁻¹ ·min ⁻¹	-
Von Heimberg et al.	2007	Norway	14M	Direct gas analysis	Without wasting time	7 mins	Mean 3.7 l·min ⁻¹ (Mean 44 ml·kg ⁻¹ ·min ⁻¹)	4.0 l·min ⁻¹ (50 ml·kg ⁻¹ ·min ⁻¹)
Dreger & Petersen	2007	Canada	30M / 23F	Direct gas analysis	As quickly as possible	8 mins	Mean 34.1 ml·kg ⁻¹ ·min ⁻¹	-
Elsner et al.	2008	USA	20M	Direct gas analysis	Firefighting pace	11 mins	Mean 29.1 ml·kg ⁻¹ ·min ⁻¹	-
Williams-Bell et al.	2009	USA	34M / 23F	Direct gas analysis	As quickly as possible	8-11 mins	Mean 39 ml·kg ⁻¹ ·min ⁻¹ (M) Mean 37 ml·kg ⁻¹ ·min ⁻¹ (F)	-
Perroni et al.	2010	Italy	20M	Direct gas analysis	As quickly as possible	8 mins	Mean 28-38 ml·kg ⁻¹ ·min ⁻¹	-

M = male participants; F= female participants

In the studies described there were considerable differences in the task design and duration of the assessed firefighting tasks. In a small number of these studies, researchers separated out the specific firefighting activities and investigated the demand of each of the tasks separately no matter how short the duration with 2 of the studies identified that ladder raising tasks, hose dragging and victim rescue lasted less than 1 minute (Gledhill & Jamnik, 1992a; Lemon & Hermiston, 1977). In another study, Bilzon et al. (2001) also separated the firefighting tasks out but ensured that each activity was long enough to tax the cardiorespiratory system and reach a steady state of oxygen consumption (4 minutes tasks).

All of the remaining studies combined tasks to develop realistic firefighting simulations which generally lasted between 8-11 minutes (Elsner, 2008; Perroni et al., 2010; Petersen et al., 2016; Sothmann et al., 1990; Von Heimburg et al., 2006; Williams-Bell et al., 2009) with one exception lasting just over 22 minutes for a standardised firefighter training exercise (Holmer & Gavhed, 2007). The tasks measured in these studies also identified significant differences in the physical demand. The lowest reported mean aerobic demands were seen during short duration firefighting tasks involving activities such as hoisting or setting up equipment involving either minimal activity or only upper body work (Gledhill & Jamnik, 1992a). These tasks elicited a mean aerobic demand of between 16-23 ml·kg⁻¹·min⁻¹. Bilzon et al. (2001) also reported a mean VO₂ of 23 ml·kg⁻¹·min⁻¹ during a 4-minute boundary-cooling task. However, this task was specific to shipboard firefighting in that the activity involves cooling the external walls of the ship bulkhead, which also appeared to require little physical movement.

At the other end of the range of the physical demands, 2 studies identified stair climbing with equipment as the most aerobically demanding task and another as rescuing casualties after climbing stairs with all 3 studies reporting a mean oxygen cost of 44 ml·kg⁻¹·min⁻¹ (Gledhill & Jamnik, 1992a; Holmer & Gavhed, 2007; Von Heimburg et al., 2006). All of these studies involved completing the tasks in full fire kit whilst wearing breathing apparatus and carrying additional equipment or rescuing casualties weighing up to 80kg. It is understandable that the combination of all body activity whilst wearing full fire kit and breathing apparatus, encountering elevation along with having to move additional loads elicited the highest physical demands. However, most of the studies

investigating the cardiorespiratory demands of firefighting report mean demands of around $34 \text{ ml}\cdot\text{kg}\cdot\text{min}^{-1}$, ranging between $30.5 - 39 \text{ ml}\cdot\text{kg}\cdot\text{min}^{-1}$ from tasks ranging from 8 to 11 minutes (Dreger & Petersen, 2007; Elsner, 2008; Perroni et al., 2010; Sothmann et al., 1990; Williams-Bell et al., 2009).

The pace at which firefighters were instructed to complete the simulations is also likely to have had a significant effect on the identified physical demand. The participants in these studies were broadly required to follow one of 3 pacing strategies: (1) at 'a controlled pace' ensuring that all firefighters performed the work at the same speed, (2) at a 'firefighting pace' representing an actual emergency situation or (3) or 'as fast as possible'. Only 2 of the 10 studies used a controlled pacing strategy (Bilzon et al., 2001; Lemon & Hermiston, 1977). Lemon & Hermiston in 1977 developed a pacing strategy by using the average pace of 14 firefighters completing the tasks prior to the study beginning. Study participants were then required to follow the 'average' step rate signalled by an audible buzzer sound (Lemon & Hermiston, 1977). Using Royal Navy firefighters, Bilzon et al. (2001) also used a pacing strategy using verbal instruction (speed up, slow down) to command the participants to maintain a previously determined minimum acceptable level of performance (Bilzon et al., 2001).

Four of the studies required participants to perform the task at a firefighting pace designed to simulate an emergency incident. However, the instructions ranged from a 'normal firefighting manner' (Gledhill & Jamnik, 1992a; Sothmann et al., 1990), or 'with no unnecessary waste of time' (Von Heimburg et al., 2006), with one study instructing firefighters to perform the task 'as quickly as possible, but at a pace representative of that at an actual fire scene' (Elsner, 2008). Whilst the researchers in these studies were attempting to evaluate the demands of a realistic firefighting pace there were large differences in the times (in minutes and seconds) associated with task completion: 05:30 – 13:54 (Sothmann et al., 1990); 04:49 – 09:30 (Von Heimburg et al., 2006); 09:00 – 17:00 (Elsner, 2008). The wide variation in task times suggest that a firefighting pace may not be clearly defined with firefighters self-selecting their operational pace, with the latter being predominantly determined by individual fitness levels.

The remaining studies required participants to undertake the firefighting simulation ‘as quickly as possible’ (Dreger & Petersen, 2007; Holmer & Gavhed, 2007; Perroni et al., 2010; Williams-Bell et al., 2009), which may identify the peak physical demands of completing firefighting tasks during a worst-case scenario. Interestingly, in all of the studies where the researchers went on to recommend minimum fitness standards, pacing strategies used were either using a controlled pace or using a firefighting pace, suggesting that this may be necessary to identify the reasonable and realistic physical demands, however only 1 study attempted to link the minimum fitness standard with completing the tasks to a minimum acceptable performance requirement (Bilzon et al., 2001).

2.2.4 Muscular strength and endurance requirements

Despite a number of studies highlighting the importance of strength and muscular endurance for operational firefighting (Bilzon, Scarpello, Bilzon, & Allsopp, 2002; Blacker et al., 2015; Gledhill & Jamnik, 1992b; Sothmann, Gebhardt, Baker, Castello, & Sheppard, 2004; Von Heimburg et al., 2006), this area of research has received relatively little attention compared to the cardio-respiratory demands of firefighting (Bilzon et al., 2001; Jamnik, Thomas, Shaw, et al., 2010; Lemon & Hermiston, 1977; Sothmann et al., 1990; Von Heimburg et al., 2006). This may in part be due to the evidence describing the cardiovascular risks involved in firefighting (Baur, Christophi, & Kales, 2012; Kales et al., 2007; Smith et al., 2013; Soteriades et al., 2011) but may also be related to the challenges in measuring the strength and endurance requirements during actual or simulated firefighting tasks.

In an attempt to identify the role that strength and muscular endurance play in firefighting activities, researchers have sought to establish the relationships between a range of fitness parameters with the critical and most demanding aspects of firefighting. Much of the research has tended to focus on common field tests of strength and muscular endurance performance such as sit-ups and grip strength, possibly with the purpose of being more readily usable by fire and rescue services (Bilzon et al., 2002; Lindberg, Oksa, & Malm, 2014; Michaelides, Parpa, Henry, Thompson, & Brown, 2011; Rhea, Alvar, &

Gray, 2004; Williford, Duey, Olson, Howard, & Wang, 1999), however other studies have investigated the relationship with more specific clinical laboratory tests (Sheaff et al., 2010). In comparison to the cardiorespiratory demands, very few of these studies have recommended minimum standards for safe and effective firefighting performance. Table 2.2 summarises the research investigating the relationship between firefighting performance on selected firefighting tasks and generic ability tests commonly reported in the literature. From the studies reported all but one of the studies investigated the relationship between maximal performances on the ability tests with the time taken to complete the firefighting task (Lindberg et al., 2014; Michaelides et al., 2011; Rhea et al., 2004; Williford et al., 1999). The remaining study (Bilzon et al., 2002) investigated maximal performances on the ability tests with the speed (in m/s) at which they completed the casualty carry and victim rescue tasks which is reflected in the direction of the correlation in table 2.2.

Table 2.2. Correlation matrix of selected firefighting performance and common generic ability tests.

Fitness Component	Stair climbing (with equipment)	Equipment pulling	Equipment / casualty carry	Victim rescue
Hand-grip (strength)	- 0.69 (Lindberg)**+ - 0.46 (Rhea)* -0.39 (Wilford)**	- 0.73 (Lindberg)**+ - 0.85 (Rhea)* - 0.55 (Wilford)** - 0.36 (Michaelides)**	0.71 (Bilzon)** - 0.50 (Lindberg)**+	0.71 (Bilzon)** - 0.79 (Lindberg)**+ - 0.41 (Michaelides)** - 0.68 (Rhea)* - 0.59 (Wilford)**
Squat (strength)	- 0.63 (Lindberg)*+ - 0.11 (Rhea)	- 0.57 (Lindberg)**+ - 0.48 (Rhea)* - 0.05 (Michaelides)	- 0.56 (Lindberg)**+	- 0.59 (Lindberg)**+ - 0.21 (Michaelides) - 0.30 (Rhea)
Bench press (strength)	- 0.73 (Lindberg)**+ - 0.39 (Rhea)*	- 0.85 (Lindberg)*+ - 0.80 (Rhea)* - 0.22 (Michaelides)	- 0.56 (Lindberg)**+ - 0.30 (Michaelides)*	- 0.82 (Lindberg)**+ - 0.31 (Michaelides)* - 0.65 (Rhea)*
Sit ups (muscular endurance)	- 0.56 (Lindberg)**+ - 0.41 (Wilford)**	- 0.51 (Lindberg)** + - 0.15 (Michaelides) - 0.22 (Wilford)*	0.58 (Bilzon)** - 0.47 (Lindberg)**+ - 0.52 (Michaelides)**	0.56 (Bilzon)** - 0.44 (Lindberg)**+ - 0.01 (Michaelides) - 0.22 (Wilford)*

Note: All correlation values are Pearson's product moment unless indicated by + = Spearman's rank correlation coefficient.

* = $P < 0.05$; ** = $P < 0.01$;

Results from table 2.2 indicate that generally speaking aspects of strength and muscular endurance appears to be important for firefighting activities. Hand-grip strength was significantly correlated with all of the firefighting tasks (stair climbing with equipment, equipment pulling, equipment or casualty carrying and victim rescue) with relationships ranging between $r = -0.36$ to $r = -0.85$. Indeed all but 2 of the relationships demonstrated a moderate or strong relationship between the variables (Bilzon et al., 2002; Lindberg et al., 2014; Rhea et al., 2004). However, where weak relationships were observed during stair climbing with equipment ($r = -0.39$) (Williford et al., 1999) and with equipment pulling ($r = -0.36$) (Michaelides et al., 2011), this was not supported by other studies. The strongest relationships with grip strength involved pulling equipment and dragging victims weighing up to 82 kg (Williford et al., 1999).

Upper body bench press strength was significantly correlated with performance in all of the firefighting tasks except one (equipment pulling), however this finding was also not supported by other studies (Lindberg et al., 2014; Rhea et al., 2004). Relationships with task performance ranged from $r = -0.22$ (Michaelides et al., 2011) to $r = -0.85$ (Rhea et al., 2004), however there was generally a mix of correlations from weak to strong across all firefighting tasks suggesting that maximum bench press strength is not consistently related to firefighting task performance. The relationships between maximum squat strength and firefighting task performance was also somewhat equivocal. Significant (but only moderate) relationships were found between maximal squatting ability with all of the firefighting tasks, however weak and non-significant relationships were found with the stair climbing (Rhea et al., 2004), equipment carry and victim rescue tasks (Michaelides et al., 2011). Muscular endurance sit up performance was significantly (but only moderately) correlated with all firefighting tasks, although weak and non-significant relationships were found with equipment pulling and victim rescue tasks (Michaelides et al., 2011; Rhea et al., 2004).

It is interesting to note that considering a number of strong and significant relationships found in these studies, with many of the authors highlighting the importance of these activities for firefighting and the applicability of the tests, none of these studies went on to recommend minimum performance standards for these tests. There are only 2 studies reported in the literature that have identified minimum strength and muscular

endurance standards for firefighters. In 1992, Gledhill & Jamnik recommended minimum sit up standards for male and female firefighters under the age of 30 years (36-39 & 29-32 sit ups for males and females respectively) and over the age of 30 years (30-32 and 23-25 sit ups for males and females respectively). However, these standards were based on the 60th percentile of the Canadian population taken from norm data and not directly related to firefighting task performances and as such may not be considered justifiable standards (Gledhill & Jamnik, 1992b). In 2004, Sothmanm et al. identified a minimum acceptable level of performance on a fire suppression task and compared this to performances on 4 predictor tests (arm endurance on a monarch arm cycle, a lateral arm lift exercise, dummy drag and a high rise pack carry). Whilst the authors were able to successfully use a combined score for the predictor tests to calculate successful performance on the fire suppression task, the tests did not use generic ability tests, potentially limiting their application. Additionally, the predictor tests included firefighting simulation tasks, which may require substantial time to set up and administer, also requiring space in which to conduct the test, further limiting practical application and implementation.

2.3 Establishing physical employment standards

A PES differs from the traditional cognitive or academic employment standards because the physical aspect required for the job for many physically demanding occupations (military, emergency services) has an important safety component associated with it. The purpose of a PES therefore has 2 main elements which has been described in simple terms as recruiting and keeping the right people in the job, whilst excluding those that are unsuitable (Petersen et al., 2016). This section will describe the considerations surrounding physical employment standards as well as the steps required within the development process to create robust and defensible standards. Finally, the last part of the section will discuss how the physiological data is analysed and interpreted to set the limits that identify the point between acceptable and unacceptable physical performance.

2.3.1 Legal considerations and physical employment standards

In many of the world's developed countries, physical employment standards are obliged to comply with a number of legal requirements in order to satisfy that the standards developed are appropriately set. These requirements fall into 2 main groups (1) health and safety requirements which are designed to protect those that may be put at risk by a PES and (2) fairness of opportunity in employment (Adams, 2016). The latter can encompass laws such as disability and age discrimination along with equality and human rights laws depending on the country in which they are applied. Whilst each country's laws, regulations and subsequent requirements will differ slightly, the general principles surrounding establishing a PES are the same, i.e. they should not be too low as to subject employees to excessive physical strain, or put others at risk from an inability to do the job. However, the standards should also not be set unreasonably high so as to unnecessarily exclude individuals from working. In essence, the standard must be based on employees being able to perform essential components of the job, safely and effectively (Jamnik et al., 2013).

2.3.1.1 Health and safety considerations

The responsibilities on employers to provide a safe place for employees to work is not a new concept, with workplace health and safety laws being established in the 1800's in both the US and the UK (Great Britain Parliament, 1974). Indeed, the concept that employees also have a duty of care to protect their co-workers and those around them whilst at work is also not novel. As firefighting with it being recognised as one of the most dangerous jobs in Britain (Roberts, 2002), some may incorrectly believe that these rules do not apply, or to the same standards for such high risk industries. Whilst the term 'safe' in the fire and rescue service is incorporated into the context of the work and 'within reason' or 'within reasonable practicability' does offer a degree of flexibility, it could be argued that workplace health and safety regulations are more multifaceted for the fire and rescue service and for other physically demanding industries.

With this in mind, there is a recognition (in both the UK and Canada) that workplace safety and the duty of care also extends to the maintenance of physical fitness

standards. It is now recognised that the employee has a duty of care to maintain role related fitness levels to carry out their duties appropriately. Where an employee fails to meet the fitness requirement (or, for example fails to co-operate by refusing to complete a fitness test), as set out in the policies and procedures that are designed to protect his / her welfare (e.g. a fitness policy), then this may constitute a breach of health & safety regulations (Great Britain Parliament, 1974; Jamnik et al., 2013). The employee may be liable to disciplinary action and / or prosecution for gross negligence manslaughter in the event of a serious accident (The Chief Fire Officers Association, 2012).

There is also an acknowledgment that the employer also has a duty of care to ensure that employees possess the physical attributes to undertake the work effectively and to avoid foreseeable risks in this regard (Blacker, Wilkinson, Bilzon, & Rayson, 2008; Jamnik et al., 2013). This should be enforced by undertaking regular or routine fitness assessments on incumbents. Where an employer fails to administer routine fitness reviews, or fails to act appropriately on information from a fitness review to ensure the safety of an employee, this may constitute a breach of health & safety regulations (Great Britain Parliament, 1974; Jamnik et al., 2013). Where death or injury occurs due to breach of the health & safety at work regulations, then this may constitute a gross breach and the employer may be liable to prosecution for corporate manslaughter (The Chief Fire Officers Association, 2012).

2.3.1.2 Equality and fairness considerations

In comparison to workplace health and safety law, anti-discrimination law in the workplace is a relatively recent development. The recent changes to international legislation in relation particularly to age, gender and disability as well as other protected characteristics under equality regulations (Great Britain Parliament, 2010) has forced organisations involved in physically demanding work to review their employment standards and ensure that are both scientifically justifiable and fair (Constable & Palmer, 2000; Payne & Harvey, 2010; Tipton, Milligan, & Reilly, 2012). Whilst many physical employment standards were historically established in good faith with the attempt to protect employees and to ensure that those particularly in public safety roles were able to carry out their roles efficiently to help protect lives, many of these standards may not

have been fairly implemented or used poor evidence in their development (Adams, 2016). Whilst a greater understanding of best practice in the field of PES development it is now important that organisations and researchers make every attempt to use these recognised methodological guidelines to develop justifiable standards and systems to embed tests and pragmatic solutions within organisations as quickly as possible.

Whilst the health and safety risks are often the main driver for the introduction of a PES, it is often legal challenges to these standards, which are the catalyst for refining or changing these employment standards. One of the most reported cases of legal challenge occurred in 1999 and was in relation to Tawney Meirion, a female firefighter for the British Columbia Ministry of Forests in Canada. Ms Meirion had failed to complete one of the mandatory fitness tests, which was a 2.5 km run in the allotted time of 11 minutes. However she considered herself fit to undertake her job as a firefighter. The court concluded that the research that the physical tests were based on was insufficient as the minimum performance standards set were based on the average performances of incumbents, which failed to distinguish between male and females. Therefore the research did not demonstrate that the performance standards were based on the minimum acceptable performance requirements for safe and efficient job performance (Jamnik et al., 2013). The court ruled that despite Ms Merion's failure to meet the fitness standard, she did not pose a safety risk to herself, her colleagues or the public. The outcome of this legal case changed the way in which PES in Canada were developed. In 1999, the Supreme Court established a set of test questions to be used in order to provide clarity around PES development and to assess whether a job standard is justifiably discriminatory (Jamnik et al., 2013):

1. 'Is the standard, policy or practice discriminatory and based on a prohibited ground'?
2. 'Was the adoption of the standard, policy or practice rationally connected to the performance of the job'?
3. 'Did the employer adopt the particular standard, policy or practice in an honest and good faith belief that it was necessary to fulfill that legitimate work-related purpose'?

4. 'Is the standard, policy or practice least discriminatory and reasonably necessary to fulfill that legitimate work-related purpose such that it would be impossible to accommodate individual employees without imposing undue hardship on the employer'?

In the UK there are also examples of legal cases being brought in relation to PES. In 1997, PC Allcock, a serving police officer for the Hampshire Constabulary applied for a job in the dog handling section. At the time, the selection criteria involved completing a 2-mile multi-terrain assault course in an allotted time, which was less for male than it was for female applicants. PC Allcock successfully argued that the lower pass mark for females constituted unlawful direct discrimination. His case was upheld and the different recruitment standards for men and women were subsequently abolished. More recently in 2011, PC Diane Bamber of Greater Manchester Police failed to meet the time of 2 minutes and 45 seconds to complete a shield carry exercise. However PC Bamber successfully argued that the performance standards associated with the test, which was developed over 30 years ago was not directly linked to the job and therefore could not be justified.

Also in 1997 a female officer police officer (Ms Dougan) in the Royal Ulster Constabulary (RUC) reserve applied to join the regular RUC. As part of the selection, she was required to pass the Physical Competence Assessment (PCA). Ms Dougan failed to complete the gender free pass standard of 3 minutes and 45 seconds. Ms Dougan argued that that the PCA was not justifiable. The tribunal concluded that whilst the test was developed appropriately, the pass standard set for the PCA circuit was not set appropriately and as such the test standard resulted in unlawful indirect discrimination. The RUC subsequently reviewed and changed the pass standard for the circuit and it has not been challenged again in an Employment tribunal.

Ultimately the setting of a justifiable and fair PES comes down to a balance between employee safety and employee rights and it is important to be able to understand the difference between direct and indirect discrimination:

- Direct discrimination occurs when someone is treated less favourably than someone else because of a certain reason (such as their race or sexual orientation).
- Indirect discrimination occurs when a practice, policy or rule is applied to everyone, affects some people more than others (such as their age or gender).

It is important to note that whilst it is illegal to directly discriminate against a protected characteristic (such as gender), it may be reasonable to indirectly discriminate as long as the employer can justify the reasoning for the employment standard (i.e. for health and safety reasons). Establishing sound PES development procedures will not only improve the likelihood of developing a fair and balanced fitness standard, it will also focus the organisation on employee safety, developing novel ways of working, and seeking technological advancements that reduce the strain on employees which will in turn help to improve the diversity of the workforce.

2.3.2 Procedures for establishing physical employment standards

The development of the theoretical concepts of PES has evolved in many ways from educational, language, psychological and health assessment settings (Zumbo, 2016) with various procedures being suggested and adopted (Blacker et al., 2015; Constable & Palmer, 2000; Payne & Harvey, 2010; Rayson, 2000a). Until recently however, there has been a lack of international consensus in the stages required in the development of robust PES. Following the first international conference on physical employment standards in 2012 (Milligan & Tipton, 2013) a number of stages were agreed as being critical to the development of robust PES (Tipton et al., 2012). These included:

1. Establish the critical tasks [task analysis]: identify the critical physically demanding tasks through task analysis - determine the number and nature of tasks to be included
2. Determine the 'Method of Best Practice' for undertaking the critical tasks
3. Agree on an acceptable minimum level of performance on the critical tasks
4. Collect physical and physiological data [physical demands analysis]: establish the

demands associated with the critical tasks and decide upon the most appropriate descriptive statistical measure (e.g. the minimum, maximum, average, percentile, mode, median etc.) to optimise employability, without sacrificing the ability to perform the critical task

5. Determine a reasonable maximum permissible relative workload e.g. the percentage of an individual's maximum work capability it is reasonable to expect them to work at
6. Production of a minimum occupational fitness standard

Two of these stages of development have been used successfully for many years, which are the task analysis and physical demands analysis.

2.3.2.1 Task analysis

The process of methodically identifying the specific activities undertaken by workers is often termed a job or task analysis and is the foundation of accurately describing and understanding an incumbents role (Constable & Palmer, 2000). Job analysis can take a number of forms, including direct observation of employees undertaking their role and systematically recording the actions of workers over a set period of time. Employee questionnaires can also be used where workers highlight the tasks they undertake and describe the frequency of completing such tasks. Other forms of job analysis include interviewing subject matter experts individually or as part of a group or technical panel to gain a consensus of the types of activities and incidents they attend.

Whilst the task analysis is used to accurately identify the critical and most physically demanding tasks in demanding occupations such as the fire service, very few of the studies investigating the physical demands of this work often fail to refer a task analysis process describing how the arduous tasks have been chosen for the research (Elsner, 2008; Holmer & Gavhed, 2007; Perroni et al., 2010; Von Heimbürg et al., 2006). Studies that have used robust task analysis processes to accurately detail the critical and most physically demanding aspect of firefighting have generally been projects involved in establishing physical employment standards suggesting that this level of scientific rigour is necessary to satisfy the organisational and legal requirements of establishing an

employment standard (Bilzon et al., 2001; Gledhill & Jamnik, 1992a; Sothmann et al., 1990; Williams-Bell et al., 2009).

In the context of PES, the job or task analysis is a critical component in the development process as it clearly identifies the steps taken and the methods used to establish the job tasks. For this reason it has been described as ‘the cornerstone for the development of job related fitness standards’ (Rayson, 2000). However, establishing these tasks is often a protracted and complex process which can be time consuming and costly to undertake (Rayson, 2000a). One of the first steps within the task analysis process is to identify the most commonly performed and physically demanding job tasks as well as those that are considered critical to the role no matter how infrequently they are undertaken (Payne & Harvey, 2010). For example, a firefighter may only be required to rescue casualties from a fire a handful of times in their career, however, this ability to be able to complete this activity successfully is critical to the safety of the public and as such should form part of the demanding tasks and subsequently be reflected in the performance standard. Establishing the method of best practice is also an important stage in the development process to review and potentially identify new ways in which to reduce the physical strain placed on incumbents. For example, over time technological advancements may change the nature of the job making equipment lighter or negating the need for physical work through mechanisation. The process should be able to clearly identify physical tasks that cannot be undertaken by other means, and for these tasks to clearly identify the most efficient way of performing them. A failure to utilise new technology (considering financial practicalities) or to recognise new ways of working could be seen as a failure of the employer’s duty to accommodate new developments in working practices (Jamnik et al., 2013).

Once the critical and most physically demanding tasks and the ways in which they should be performed have been established, clarifying and identifying performance standards for these tasks is a vital step in the task analysis process. This may include for example 3 categories; “acceptable”, “minimally acceptable” and “unacceptable” performances or 2 categories with “minimally acceptable” and “unacceptable” performance. The critical part here is the identification and description of the minimally

acceptable standard as this will translate into the pass / fail cut points on the performance test (Zumbo, 2016).

Considering the importance of this phase, there is a lack of consensus on the best practice methods of establishing these requirements, along with a lack of published task analyses that clearly define the steps involved in determining these minimum acceptable performance requirements. A number of studies have used a representative sample of employees to perform tasks at self-selected work rates and assumed current employees would be capable of performing the task at a reasonable pace (Jamnik, Thomas, Burr, & Gledhill, 2010). However in emergency service roles, in particular the fire and rescue services, the number of emergency calls that UK firefighters attend has dropped sharply in recent years (Office of the Deputy Prime Minister, 2003) and previous reports have identified that the physical demands of the job appear to be insufficient to maintain role specific fitness levels (Office of the Deputy Prime Minister, 2004b). It may therefore be wrong to assume that the self-selected pace is appropriate for this group or workers, as many firefighters may not have performed these critical and demanding tasks for some time and with enough regularity to ensure a consistent and appropriate task performance (Zumbo, 2016).

Conversely, other researchers have raised concerns that employees being placed in a test type situation may feel pressured to work harder and / or faster than they normally would (Hawthorne effect) purely because they are being scrutinised (Tipton et al., 2012). This would have the effect of making a task appear more demanding than it actually may be, potentially causing standards to be set unreasonably high. Indeed, recent case law has highlighted the importance of establishing standards on the minimum acceptable performance requirements rather than on average performances by workers (Jamnik et al., 2013; Tipton et al., 2012).

2.3.2.3 Physical demands analysis

The physical demands analysis is the process of quantifying the physical and physiological demands of the critical tasks (Milligan, Reilly, Zumbo, & Tipton, 2016; Tipton et al., 2012). Studies have tended to focus on quantifying metabolic and

cardiovascular strain through the measurement or estimation of physiological responses such as oxygen uptake (VO_2), blood lactate accumulation or heart rate responses (Constable & Palmer, 2000). Many of the methodological considerations related to the physical demands analysis process relate to the scientific principles of accurate measurement of the physiological responses, as well as ensuring the validity and reliability of the equipment used (Milligan et al., 2016).

However, arguably one of the most challenging areas of this process is not scientific or technical, but practical in terms of choosing and ultimately ensuring a representative sample of study participants. Without a typical representation of the workforce, it is unlikely that any performance standards will be accepted by the employer or wider stakeholder community. In physiological research, physical characteristics such as age, gender and body composition may be considered the most critical, however other characteristics such as ethnicity may also be important. As only 4% of firefighters in the UK are female (Office of the Deputy Prime Minister, 2003) many of the studies (e.g. cardiorespiratory demands research) often site data only for male participants (Elsner, 2008; Holmer & Gavhed, 2007; Lemon & Hermiston, 1977; Perroni et al., 2010; Sothmann et al., 1990; Von Heimburg et al., 2006), with average ages of participants often being in their 30's or younger (Elsner, 2008; Holmer & Gavhed, 2007; Lemon & Hermiston, 1977; Perroni et al., 2010; Von Heimburg et al., 2006) and with mean body composition values in the lean classification (Elsner, 2008; Von Heimburg et al., 2006), which may not be representative of the true working population (Baur, Christophi, Tsismenakis, Jahnke, & Kales, 2012; Goheer, Bailey, Gittelsohn, & Pollack, 2013; Munir, Clemes, Houdmont, & Randall, 2012). When presenting the descriptive physical characteristics of their population, the vast majority of published studies unfortunately have failed to demonstrate the representativeness of their target population.

Following the data collection period, it is important to agree on two important factors, which are: (1) what is the minimum demand of the task or group of tasks? (2) What is the maximal permissible relative workload for the task(s)? Because of the natural inter-subject variability in physiological data, the decision of where the minimum demand should be applied within the spread of data is an important aspect of the data interpretation process. It is important that this decision is justifiable as it will undoubtedly have

implications to the final pass score set for the PES (Tipton et al., 2012). Previous researchers have chosen the mean value, insisting that this is the point that an untested individual will most likely be closest to, assuming a normally distributed data set (Bilzon et al., 2001; Tipton et al., 2012).

However other researchers have used +2 standard deviations above the mean on timed firefighting simulations for U.S. firefighters (Sothmann et al., 1990). Canadian researchers have also used -1 standard deviations below the mean for an aerobic requirement and +1 standard deviations for timed task simulations (Jamnik, Thomas, Shaw, et al., 2010) thus incorporating 83.3% of the performances within the employment standard. It is not clear within the literature on the justification on this chosen position, however Canadian regulations in the development of a *bone fida* occupational fitness standard stipulate that cut-point must be statistically derived and also need to be above the '80% rule' in order to satisfy adverse impact regulations. Importantly, this has been upheld in human rights challenges in Canada (Jamnik et al., 2013). However, it should be noted that, while such approaches can be inclusive, many participants will be considered to have passed the test, when they would have actually failed (i.e. false positive). This may expose them to occupational tasks, which they do not have the physical competencies to complete safely.

Possibly the last decision in this process involves determining what percentage of an individual's maximum capacity could they reasonably be expected to sustain on a given task (Tipton et al., 2012). This fraction will however depend on the precise nature and duration of the task, but this information should have been defined in the task analysis process. It would clearly not be reasonable or indeed safe to expect employees to work for a particular duration above an intensity that an individual can realistically accommodate as this could lead to the increased risk of injury and accidents (Epstein, Yanovich, Moran, & Heled, 2013). This area of consideration is rarely discussed within the scientific literature, yet its implications are of vital importance to the safety of employees. Indeed much of the pioneering work undertaken appears to have not been progressed in recent years (Bink, 1962; Louhevaara, Smolander, Korhonen, & Tuomi, 1986). A number of studies that have referred to this research to determine a maximal permissible workload and ultimately a PES. In 1992, Gledhill & Jamnik, identified that a

work capacity of 85% of VO_{2max} was achievable which was based on work by Astrand and Rodahl in 1970 (Astrand & Rodahl, 1970). However, it has been identified that this evidence was based on a single subjects response to high intensity exercise (Tipton et al., 2012), thus representing a relative weak evidence base.

In 2001 Bilzon et al. proposed a work rate of 80% of VO_{2max} during shipboard firefighting citing the work of Bink et al. (1962) and Louhevaara et al. (1986). Bink et al. (1962) made a number of calculations of the maximal physical work capacity of a ‘normal’ 35 year old male for 4 minutes, 8 hours and 24 periods of continuous work (Bink, 1962). Louhevaara et al. (1986) investigated the effects of wearing breathing apparatus during light, moderate and heavy exercise on a treadmill using 13 male firefighters as subjects concluding that wearing the firefighting equipment decreased maximal working times by over 25% (Louhevaara et al., 1986). It appears quite clear that for such an important point in establishing a PES the evidence in this area requires further attention, with more participant numbers, including female participants. The added burden and additional physiological effects of wearing a range of PPE during direct task simulations should also be considered.

2.3.3 Physical employment tests

2.3.3.1 Types of physical employment tests

It has been suggested that physical employment tests (PET) can be categorised into 3 types of test (Payne & Harvey, 2010):

1. Generic predictive tests (GPTs)
2. Task simulation tests (TSTs)
3. Task related predictive tests (TPTs)

GPTs are common non-specific tests that measure a particular physical characteristic (for example cardiorespiratory fitness), which can be used in a broad range of occupational settings to predict performance on a job task. These tests are often established physical fitness tests such as an aerobic capacity step test (Buckley, Sim, Eston, Hession, & Fox, 2004). TSTs are actual replications of either a single job task or

a number of job tasks completed in sequence to represent a compilation of the critical and / or most arduous job tasks. TPTs fall somewhere in the middle of these two other test types where they are not an actual job task but are also not generic in that they include some of the characteristics of the job task (e.g. a shoulder press test to simulate lifting a fire service ladder overhead).

As the main purpose of any PET is to correctly identify employees that can and cannot do the job, it would seem a straightforward decision to opt for the test that gave the highest predictive ability. However, in a workplace setting, a vast number of practical considerations exist including equipment and resource availability, safety issues, and financial constraints making the choice of test as much about one that is able to be applied within the occupational setting as the accuracy of the test itself. However, whilst these practical considerations need to be considered, so do the concepts of test reliability and validity to ensure that the testing process correctly identifies those that can do the job as fit and those that cannot as unfit.

2.3.3.2 Reliability and validity of physical employment tests

The concepts of validity and reliability have been described as being interconnected in this field of study in that a PES or PET cannot be considered valid if it is not reliable (Milligan et al., 2016). In simple terms validity is the degree to which a test measures what it is supposed to measure and reliability refers to the test being able to produce consistent scores either at different times on the same group of individuals or when assessed by different test administrators on the same occasion (Payne & Harvey, 2010). There are 3 main categories of validity:

- Content validity (sometimes described as logical or face validity): A test has content validity if it accurately represents the components of the job task such as in a direct task simulation or TPT. Content validity is often favoured by employees and by the courts as it is seen as being more ‘job related’ (Milligan et al., 2016)
- Criterion validity (includes both predictive and concurrent validity): A test is said to have criterion validity when an association has been proven between a

test (such as a GPT or a TPT) and a direct measure of job performance (predictive validity) or with an indirect measure of job performance (concurrent validity) (Payne & Harvey, 2010)

- Construct validity (which includes convergent validity and discriminant validity, which must both be present to be considered valid): A test demonstrates construct validity when it determines the relationship between a test and a job task and when it is capable of separating employees that are able and not able of completing the critical tasks to the minimum acceptable performance standard (Milligan et al., 2016)

The reliability of a test relates to the measure of consistency within the data and refers to a number of situations:

- Test-retest reliability measures the reproducibility of test performance when undertaken on the same group of subjects at different times and can include the variability in measurement error in testing equipment such as a metabolic analyser
- Inter-rater reliability measures the reproducibility of test performance when assessing the same group of subjects with different test administrators
- Intra-subject variability measures the biological variability involved in physical performances after controlling the test for factors including familiarisation and weather conditions

There is often a desire by those involved in PES and PET to define acceptable job performance in binary terms (pass / fail). However, it is important to recognise that despite the simplicity of a cut score, predictive tests are rarely 100% accurate and therefore misclassification is possible and, both statistically and practically, highly probable. Through an understanding of the validity and reliability of the test protocols an organisation or test administrator can consider test scores close to the cut score and to be categorised as 'borderline'. In this case an organisation may choose to initiate further assessment or consider other factors that may alter the decision in relation to the implications of a fitness test score.

2.3.4 Setting standards

The concept of establishing physical standards for demanding occupations is not a new development with basic anthropometric criteria being used since Roman times to restrict physically unsuitable persons from serving in the military (East, 2013; Stout, 1921). The process of applying physical standards to military personnel remains today (Rayson et al., 2000) with the practice spreading to a number of other physically demanding jobs where the physical status of employees is considered critical to the job (Jamnik et al., 2013; Reilly, Iggleden, Gennser, & Tipton, 2006; Reilly, Wooler, & Tipton, 2006). Whilst there have been many changes to the way in which standards for these demanding jobs are established and applied, the rationale for setting them has remained the same i.e. that employees with a physical capacity below a particular point on a physical test scale will put job success and / or employees or the public at risk, therefore the physical standards are deemed necessary to undertake the job effectively. The defensibility of such physical performance standards ultimately relies on the ability to clearly demonstrate that the physical attribute or test is linked to job performance and that a proposed passing score on the test corresponds to an appropriate performance standard (Kane, 1994). It is this process which is commonly referred to as standard-setting (Zumbo, 2016).

2.3.4.1 Identifying performance standards

A performance standard has been defined as the conceptual version of a desired level of competence (Kane, 1994). Being able to clearly define this desired level of competence within the research process ensures that there is a clear understanding about what performance is expected or required. Whilst traditionally the terms “acceptable” and “unacceptable” may have been used to delineate job performances, in the field of PES however, the term “minimal acceptable performance” is commonly used as this is the point where an employer may more easily be able to justify that they have attempted to balance both the health and safety aspects of the work with the equality considerations of its employees. Where more than two performance standards are identified, it may also be important to describe “acceptable” job performance and not just the minimum requirement. A lack of clarity between “acceptable” and “minimally acceptable”

performance standards could impact on the pass score applied to a test which may subsequently unfairly impact on applicants or incumbents, particularly those in physically disadvantaged groups such as females and older incumbents (Adams, 2016).

2.3.4.2 *Establishing cut scores*

A cut-score (or pass score) had been described as the ‘operational version of the desired level of competence’ which is a ‘distinct point on the test score scale’ (Kane, 1994). Numerous methods have been developed for establishing cut scores (Kane, 1994), however within PES development cut-scores are generally determined by one of 2 methods (Zumbo, 2016):

1. On the statistical distribution of test scores
2. From the judgements of subject matter experts

The use of statistically derived cut-scores (e.g. using the mean time +1 standard deviation to complete a work simulation) are free from human intervention and potential bias and have successfully been used to implement and more importantly defend employment standards in physically demanding occupations (Jamnik et al., 2013). However, they are often described as arbitrary because the cut-point is set purely by statistical means and without any other evidence or data, there is not a compelling argument why it could not easily be set somewhere else on the test score scale (Zumbo, 2016). On the other hand, cut scores which are determined from the judgements of subject matter experts (e.g. training staff) are not determined statistically but based on a group consensus of experiential knowledge. However, human and methodological factors can potentially influence the cut-score chosen which may include the design of the voting system, the number, diversity and experience of the experts used in the panel, which can all introduce variability in the standard-setting process. It is recognised that because there are a range of legitimate choices that can be made with any of the methods used to determine a test passing score, one must accept that there is a degree of arbitrariness in all cut-score and standard-setting (Kane, 1994; Zumbo, 2016). However, this is not to say that standard-setting *per se* is random as standards can vary in their arbitrariness (Kane, 1994).

2.3.4.3 Evaluating performance standards

Often the final stage in best practice PES development is to evaluate the effectiveness of the employment tests and associated passing scores (Tipton et al., 2012). The use of a decision theory matrix (or contingency tables) is one of the most common methods used (Constable & Palmer, 2000; Petersen et al., 2016). Contingency tables (Figure 2.1) with their associated calculations can be used to identify the number of true and false positives (i.e. the number of people that pass an employment test, which can or cannot do the job) as well as number of true and false negatives (i.e. the number of people that fail an employment test that can and cannot do the job).

Following this, other calculations such as test sensitivity and specificity as well as positive and negative predictive values can be used to help improve test validity by helping to identify the most suitable cut-score and ultimately to help justify its position on the test score scale.

		The "Truth"	
		Yes	No
Test Result	Yes	(A) True +	(B) False +
	No	(C) False –	(D) True –

Figure 2.1 Decision theory matrix (contingency table)

In the context of PES development:

- ‘Sensitivity’ (often called the true positive rate) identifies the percentage of individuals who can do the job that pass the test and is calculated by: $\text{True Positives} / (\text{True Positives} + \text{False Negatives}) \times 100$
- ‘Specificity’ (often called the true negative rate) identifies the percentage of people who cannot do the job that fail the test and is calculated by: $\text{True Negatives} / (\text{True Negatives} + \text{False Positives}) \times 100$
- The ‘positive predictive value’ identifies the percentage of individuals which pass the test can actually do the job as is calculated by: $\text{True Positives} / (\text{True Positives} + \text{False Positives}) \times 100$.
- The ‘negative predictive value’ identifies the percentage of individuals that fail the test cannot actually do the job as is calculated by: $\text{True Negatives} / (\text{True Negatives} + \text{False Negatives}) \times 100$

Whilst these calculations are revealing when attempting to establish an appropriate cut score, one also has to accept that within this process other factors such as the imperfections of tests, the trade-off between sensitivity and specificity (i.e. where sensitivity increases, specificity decreases and *vice versa*), as well as the need to be cognisant of the financial and / or resource implications of test development may all play an important part in the final decision about where to set a cut score. In relation to specificity and sensitivity, it is important that employers consider the implications of false negatives and false positives within the context of the work that the test is to be used. For example an employment test used in the emergency services may wish to minimise the number of false positives, thus minimising the risk of someone being in a safety critical role without the required levels of physical fitness with the potential of placing the public or their colleagues at risk from substandard job performance. However, this will consequently increase the number of false positives (the proportion of incumbents failing a fitness test that can do the job). In this case other measures such as further testing methods may be necessary to improve worker protection in these physically demanding jobs.

Because of the many influences in the workplace in relation to the development of an employment test cut-score, the standard-setting process has been described as a 'policy decision' (Zumbo, 2016). It is therefore important that researchers and employers describing in detail its procedures and decisions when establishing employment tests whilst at the same time following best practice guidance. In turn, this should reduce the degree of perceived or actual arbitrariness in the standards setting process with the intent to increase the defensibility of the physical employment standard being developed.

CHAPTER 3

Study 1

A Task Analysis Methodology for the Development of Physical Employment Standards

3.1 Introduction

Workers that perform public safety occupations undertake a variety of activities that can be both hazardous and physically demanding (Jamnik et al., 2013; Roberts, 2002). These individuals are often required to respond within minutes, transitioning from rest and occasionally sleep, to high levels of physical exertion (Bos et al., 2004). Consequently, a number of international studies have identified the importance of physical fitness in public safety roles (Gumieniak, Jamnik, & Gledhill, 2013) and subsequently quantified the physical and/or metabolic demands of strenuous safety-related occupations, including: correctional officers (Jamnik, Thomas, Burr, et al., 2010), police officers (Jamnik et al., 2013), ambulance service workers (Barnekow-Bergkvist, Aasa, Angquist, & Johansson, 2004; Gamble et al., 1991), military personnel (Wilkinson, Rayson, & Bilzon, 2008) and firefighters (Bilzon et al., 2001; Elsner, 2008; Holmer & Gavhed, 2007; Von Heimburg et al., 2006).

Two key stages often used in the process of determining the physical fitness requirements for a safety-related occupation are: (i) a task analysis and; (ii) a physical demands analysis. The aim of a task analysis, particularly when determining minimum occupational fitness requirements, is to clearly identify the critical and most physically arduous generic aspects of a job (Rayson, 2000; Taylor & Groeller, 2003; Truxillo, 2004) and to determine the minimum acceptable performance requirements. A physical demands analysis would then typically follow, and would involve the collection of physiological and/or physical performance data to quantify the physical demands of the tasks identified in the task analysis, performed to the minimum standard (Constable & Palmer, 2000; Payne & Harvey, 2010). Whilst many task analyses precede physical demands analyses, few have articulated the practical steps taken in a systematic manner in order that they could be replicated in other settings (Blacker et al., 2015). Additionally, a limited number of task analyses have been completed with the specific foresight to inform a future study aiming to quantify the physical demands of, and therefore the physical requirements for, tasks performed to a “minimum acceptable” requirement

(Holmer & Gavhed, 2007; Von Heimburg et al., 2006). Ultimately, it is upon these requirements that minimum fitness standards should be based. Finally, the interim process of developing representative simulations of physically arduous tasks and objectively determining what constitutes minimum acceptable performance is also pivotal in ensuring the acceptability and validity of resultant standards, both to employees and employers.

In a number of developed countries, the implementation of justifiable physical employment standards for arduous jobs has become increasingly important. Changes to legislation around discrimination, in particular on the grounds of disability, age and sex has highlighted the legal requirement to develop fair and unbiased physical fitness standards (Jamnik, Gumienak, & Gledhill, 2010; Jamnik et al., 2013; Payne & Harvey, 2010; Tipton et al., 2012). In addition, ensuring that employees maintain appropriate levels of physical competence, by administering routine physical fitness tests, is also now recognised as an important part of an employer's on-going 'duty of care' to help safeguard the health and safety of their employees (Great Britain Parliament, 1974; Jamnik et al., 2013). It is therefore important that both pre-employment and incumbent fitness standards be based on the physical demands of the tasks, which employees are expected to perform.

In the UK fire & rescue services, previous work to determine critical and arduous tasks has been undertaken for point-of-entry, or pre-employment, testing (Blacker et al., 2015; Rayson, Wilkinson, Carter, & Nevill). However, the metabolic and cardiovascular demands of tasks performed by serving firefighters to a minimum acceptable requirement have not been quantified, which has hindered the development of evidence-based fitness standards for incumbents. Indeed, it is not possible to conduct a physical demands analysis without having first conducted a systematic task analysis, which provides sufficient information to subsequently determine minimum occupational fitness standards. Whilst frameworks of the key stages for developing occupational fitness standards have been published (Tipton et al., 2012), the practical steps required to fulfil these frameworks are not often documented. A proposed model for such a systematic task analysis process appears to be lacking from the published literature. To our knowledge, this will be the first paper to describe and document a practical model of a structured task analysis process used to, specifically, define and agree the minimum acceptable performance standards of essential generic occupational tasks. This process is essential

for informing the development of minimum occupational fitness standards for a physically demanding occupation.

3.2 Methods

A task analysis of the critical and most arduous generic firefighting tasks was undertaken in the UK fire & rescue service between October 2012 and March 2014. The research team collaborated with key stakeholders from the Chief Fire Officers Association (CFOA). We followed a framework of principles identified previously (Tipton et al., 2012), which included the following key stages:

1. Establish the critical tasks
2. Determine the “method of best practice” for undertaking the critical tasks
3. Agree on an acceptable minimum level of performance on the critical tasks

This study attempted to expand on these key stages by detailing the practical steps required within a task analysis process needed to satisfy industry stakeholders in the development of an occupational fitness standard for a physically demanding occupation.

Project working groups

Two distinct working groups of subject-matter experts were established to provide the research team with, technical and strategic review and guidance relating to the job (e.g. UK firefighting). A Technical Panel (TP), consisting of operational personnel, was assembled to advise on the practical aspects of the job, whilst a Stakeholder Panel (SP) was established to provide strategic direction to the project team, to ensure that the process and outcomes were both logical and justifiable to the customer. Whilst the SP did not affect decisions made by the TP, they did evaluate and finally endorse all major decisions. The two panels were kept independent from one another throughout the project to ensure that political and/or strategic motivations did not influence alternative group outcomes, whilst the research team facilitated the transfer of information between the groups.

Technical panel (TP)

The TP consisted of 13 male operational personnel aged (mean \pm SD) 41 ± 7 years, from 10 fire and rescue services across the UK, with a range of ranks (e.g. firefighters, crew and station managers) and an average of 17 years of experience (range 10-27 years). Panel members were nominated from national technical working groups and were selected on their expertise and recent experience in operational incident management or in the delivery of training in one or more of the following areas; equipment manipulations (water relays using fire service hose / ladders / portable pumps); the use of breathing apparatus in structural fires; incidents involving chemical protection suits, wild-land firefighting, rope rescue, water or mud rescue, road traffic collisions and urban search & rescue activities. While a sex-diverse panel would have been preferable, unfortunately no female personnel volunteered to participate on the panel.

Stakeholder panel (SP)

The SP consisted of nine (8 male, 1 female) senior managers (i.e. Chief and Deputy-Chief Fire Officers) from UK fire & rescue services (age range 45-60 years) leading national working groups on firefighter fitness, health & safety, occupational health and technical response. The panel also embraced representation from the trade unions and local government association.

Task analysis process

A series of focus group meetings were conducted by the research team, which consisted of the TP examining relevant literature (Office of the Deputy Prime Minister, 2004a; Rayson et al.) and fire service documents (Her Majesty's Fire Service Inspectorate, 2004) reviewing best practice methods and discussing experiences within the group in open discussion before reaching a group consensus on any decisions required for the research process. This guaranteed that all decisions relating to the technical aspects of firefighting were made independently, by the subject-matter experts. These collective TP decisions were then taken to the SP for endorsement before moving on to each subsequent phase of the project (Figure 3.1).

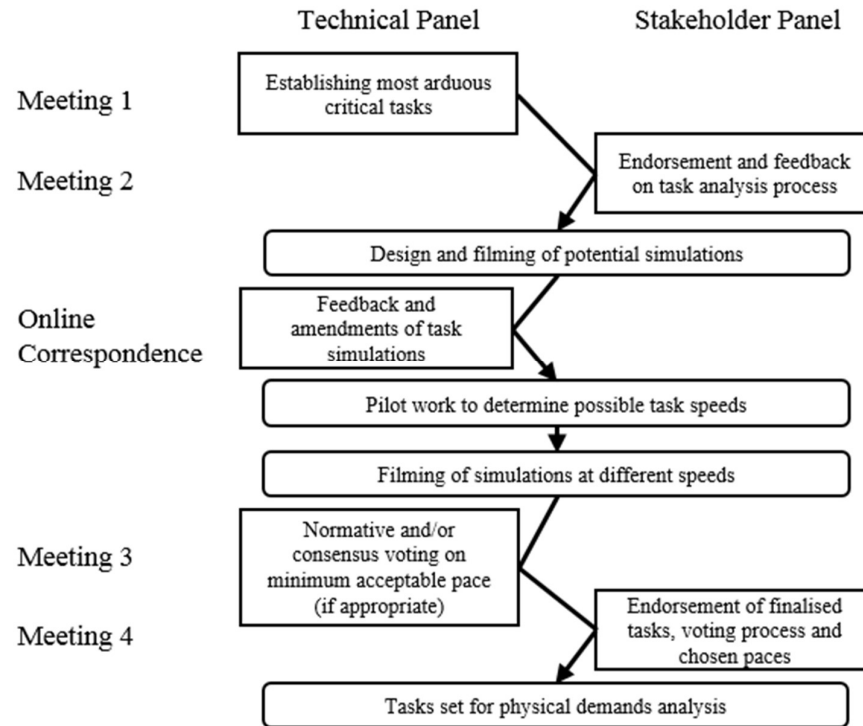


Figure 3.1. Task analysis consultation process using convened meetings with technical (subject-matter experts) and stakeholder panels.

Establish the critical tasks

Several meetings were convened for the TP to identify, discuss and agree upon the critical and most physically demanding aspects of UK fire and rescue activities. Initially, the TP were tasked with identifying any specific role-related differences within the rank structure of UK fire service personnel. Consideration was also given to whether any other factors (such as age and sex) would alter job role. In the UK fire & rescue service, any operational firefighter is expected to complete the same tasks irrespective of age and sex. Following this, activities that were considered to be specialist roles (including road traffic collisions or water rescue activities) were dismissed from subsequent analysis on the basis that they would not be generic to all firefighters. Only tasks that were deemed to be both critical and the most physically demanding for all UK firefighters were included.

Determining the method of best practice

The TP identified the safest, most efficient manner of performing each of the tasks while adhering to established training guidelines, standard operating procedures (Her Majesty's Fire Service Inspectorate, 2004) and safety regulations (Health & Safety Executive, 2004). In order to assist in determination of minimum acceptable performance of tasks, realistic simulations were developed by the TP to reflect the role of one individual in activities that incumbents would reasonably be expected to perform as part of their operational role. Typical distances and equipment used were agreed upon by panel members. The simulations were designed to fulfil the following criteria: being easily replicable (i.e. reproducible on a fire service training ground using standard fire service equipment); easily regulated (in terms of pace and instruction). With the specific foresight that a task analysis is often used to inform a subsequent physical demands analysis, it was also considered (if applicable) that tasks (while not measured in this study) should be of sufficient duration to elicit a representative steady-state of oxygen demand (for use in a future physical demands analysis). Finally, to attempt to establish the “urgency” around each task for when it would be performed, a hypothetical occupational scenario was constructed to provide specific context for that task.

Agree on an acceptable minimum level of performance

Once the task simulation protocols had been agreed upon, the appropriate simulations were filmed being performed by a trained male incumbent at three varying paces (video A – “slow” pace, video B – “moderate” pace, and video C – “fast” pace). The “moderate” pace (video B) corresponded to the average pace of two training instructors performing the task(s) at self-selected pace typical of an emergency incident. The slower and faster paces were chosen by adding (or subtracting) round increments of speed to the moderate pace while being both a) visually dissimilar from the moderate pace for easy differentiation and b) still within a safe pace for the nature of the simulated task(s).

The pace of the trained male incumbent performing each of the tasks was kept constant using a number of methods depending on the type of activity being performed. For activities involving walking or running over ground, the pace was controlled by passing marker cones (placed at 5-metre intervals) in time with audible signals emitted

from an audio player. For tasks involving stair climbing and extending ladders, the pace was controlled using a metronome through headphones to indicate the appropriate step / pull rate, respectively. Five male incumbents (mean \pm SD, age: 40 ± 4 y, height: 1.77 ± 0.05 m, body mass: 83 ± 8 kg) were used for the filming of the task paces (the same individual was consistent for each task). These individuals were sought to represent the average UK firefighter (age: 42 ± 7 y, height 1.79 ± 0.07 m, body mass: 86 ± 13 kg, unpublished data) in an attempt to mitigate any visual bias to the perception of ease or difficulty of the task on film. While a sex-divergent group of incumbents who were used for the filming would have been preferable, no female incumbents volunteered to participate.

To determine the minimum acceptable level of performance for each critical task the Bookmark method of standards setting was adopted. Technical panel members were shown the videos of each simulation being performed at the three paces (in sequence from slowest to fastest) and were asked to indicate what they felt was the minimum acceptable requirement for each task. Each TP member voted anonymously on a scoring sheet for the pace that they felt corresponded to the minimum acceptable performance of the specific task (within the context of the scenario described). Panel members were given the option to choose the speed indicated by the videos shown, and also the speed between those videos, thus giving five choices in total. For some tasks, such as lifting a mass overhead, successful or unsuccessful completion was discrete (pass/fail) and therefore did not require judgement on any appropriate pace.

The actual pace of each displayed task was not divulged to the panel members so as not to influence their decision in any way. The individual votes from TP members were collated and presented back to the panel. The TP were then asked to reach a group consensus for each task. Normative analysis (mean and mode) of the votes was used to indicate the possible minimum acceptable pace, and was brought to discussion. Where responses clearly indicated a majority (mode) response, this pace was selected for discussion. Where a response was split between two choices, the middle point between the two choices was selected for discussion. Where a clear majority decision was not reached, further discussion took place around best practice of the activity and the context of the simulation until a consensus was reached for these tasks. It should be considered

that if the votes are markedly polarised among the panel and, following discussion and clarification, it is clear that a consensus cannot be agreed, the task itself should be reconsidered, altered or excluded from further consideration.

3.3 Results

The TP identified two distinct functions in UK firefighters and clear differences between operational firefighting roles and incident command roles. Those in a “firefighting” role (typically the rank of Firefighter, Crew Manager and Watch Manager) performed the most arduous of firefighting duties (casualty evacuation; equipment carrying; hose running; stair climbing; wild-land firefighting; lifting ladders; extending ladders; lowering ladders), whilst fire-ground “incident commanders” (typically the rank of Station Manager and above) were involved with reaching the operational incident (by walking and climbing stairs at wild-land fires and high-rise building fires respectively) and supervising firefighters at the operational scene. It was agreed that incident commanders would not be expected to undertake activities identified for those in a firefighting role. However, it was considered reasonable for this group of employees to wear the same personal protective equipment as a firefighter whilst reaching, and in attendance at, the operational incident.

Developing realistic simulations

Realistic single-person simulations were developed to reflect the activities that incumbents would be expected to perform as part of their role. The available choices of acceptable pace for each of these activities shown to the TP are displayed in table 3.1. Descriptions of the simulations are described below:

- *Hose run task (firefighter)* – A simulated water relay task to establish a water supply from a fire hydrant to a fire appliance 100 m apart using a total of four lengths of hose completed over a flat 25 m course

- *Casualty evacuation task (firefighter)* –A simulated entry to, and rescue of an unconscious casualty from, an industrial building whilst wearing breathing apparatus equipment
- *Equipment carry task (firefighter)* –A simulated equipment-handling task carrying firefighting equipment over a 200 m distance. Performed by walking a flat 25 m course while carrying a 25 kg barbell
- *Wild-land fire task (firefighter)* – A simulated wild-land fire suppression task over 200 m using a fire beater. Performed by traversing a 50 m course of sloped rural ground 4 times, beating the ground on each ascent
- *Wild-land fire task (incident commander)* – The simulated management involvement during a wild-land fire. Performed by walking a 50 m course of sloped rural ground 4 times (without fire beating)
- *Stair climbing task (firefighter)* – A simulated high-rise building fire. Performed by climbing 12 flights of stairs whilst wearing breathing apparatus equipment carrying 25 kg of firefighting equipment
- *Stair climbing task (incident commander)* - The simulated management involvement during a high-rise building fire. Performed by climbing 12 flights of stairs whilst wearing breathing apparatus equipment (without equipment)
- *Ladder lift task (firefighter)* – A simulated ladder lift, lifting $\frac{1}{2}$ of the weight of the head of a 13.5 m fire service ladder. Performed by lifting a bar on a pivot arm from hip height to 1.82 m overhead (Approximately 29 kg at the mid-lifting point)
- *Ladder lower task (firefighter)* – A simulated unhooking of a 13.5 m ladder in order to lower the equipment using a ladder simulator. Performed by a single overhead downward pull on a rope with both hands (Approximately 42 kg)
- *Ladder extension task (firefighter)* – A simulated extension of a 10.5 m fire service ladder using a wall-mounted ladder simulator. Performed by continuously pulling down (hand-over-hand) on a rope until full extension (Approximately 28 kg)

Table 3.1. Speeds of each recorded video, for each task and corresponding voting options.

Video	Video A		Video B		Video C
Voting options	1	2	3	4	5
Hose Run	6 km/h		8 km/h		10 km/h
Equipment Carry	4 km/h		6 km/h		8 km/h
Stair Climb	75 steps/min		95 steps/min		115 steps/min
Casualty Evac. (H)	4 km/h		8 km/h		10 km/h
Casualty Evac. (C)	2 km/h		3 km/h		4 km/h
Wild land fire	2 km/h		3 km/h		4 km/h
Ladder Extension ⁺	30 reps/min		70 reps/min		110 reps/min

⁺ *reps/min = repetitions (rope pulls) per minute; (H)=hose section; (C)= casualty evacuation section.*

The mean, mode, range and consensus for the minimum acceptable paces for each simulation are shown in table 3.2. Both the TP and SP agreed and endorsed, respectively, that each of the single-person simulations developed for the determination of the minimum acceptable pace used up-to-date best practice methods, accurately reflected reasonable expectation of a firefighter (or incident commander), and the minimum acceptable requirement for each of the tasks. Simulations that had been developed previously in other related projects (Blacker et al., 2015), which were deemed to still employ best practice were included within the battery of simulations. While a majority (mode) vote existed for task pace, the wild-land fire task was the only task to receive the full range of votes (1-5).

Table 3.2. Technical panel choices, mean, mode, range and consensus scores with corresponding minimum acceptable work rates.

Task	Vote score (Mean \pm SD)	Vote score (Mode)	Vote range	Consensus score	Chosen pace
Hose Run	2.8 \pm 0.4	3	2-3	3.0	8.0 km/h
Equipment Carry	2.3 \pm 0.9	2	1-4	2.5	5.5 km/h
Stair Climb	3.1 \pm 0.7	3	2-4	3.0	95 steps/min
Casualty Evac. (H)	3.5 \pm 0.8	3,4	2-5	3.0	6.0 km/h
Casualty Evac. (C)	3.0 \pm 0.9	3	2-5	3.0	3.0 km/h
Wild land fire	3.9 \pm 1.2	4	1-5	4.0	3.5 km/h
Ladder Extension ⁺	3.3 \pm 1.0	3	2-5	3.0	70 reps/min

⁺ *reps/min = repetitions (rope pulls) per minute; (H)=hose section; (C)= casualty evacuation section.*

The bespoke steps of the task analysis identified within this study are summarised in table 3.3.

Table 3.3. Summary of the practical steps undertaken.

Phase	Step
1. Establishing the critical tasks	<i>(a)</i> identifying the most physically demanding and critical tasks <i>(b)</i> disregarding specialist activities <i>(c)</i> identifying role related differences where necessary
2. Determining the “method of best practice”	<i>(a)</i> identifying standard operating procedures <i>(b)</i> developing realistic single–person simulations <i>(c)</i> identifying task-specific contextual scenarios
3. Agreeing on an acceptable minimum level of performance	<i>(a)</i> developing a pacing strategy <i>(b)</i> identifying an objective scoring system <i>(c)</i> gaining consensus agreement

3.4 Discussion

This study describes a task analysis designed to identify the minimum acceptable performance requirements of the critical and most physically demanding tasks within a safety-related occupation. We have expanded on the key stages identified previously (Blacker et al., 2015; Rayson et al., 2009a) by identifying bespoke steps within each stage of the task analysis process of: 1) establishing the critical tasks (identifying the most physically demanding and critical tasks; disregarding specialist activities; identifying role related differences where necessary); 2) determining the “method of best practice” (identifying standard operating procedures; developing realistic single–person simulations; identifying task-specific contextual scenarios) and; 3) agreeing on an acceptable minimum level of performance (developing a pacing strategy; identifying an objective scoring system; gaining consensus agreement).

In the present study, this was achieved through consulting with subject-matter experts and the use of single-person simulations, video analysis and the “Bookmark method” of standard setting (Lewis, 1999) to determine the minimum acceptable performance requirements of the most physically demanding and critical tasks undertaken

by UK firefighters, specifically. This was performed so that the cardiorespiratory, strength and muscular endurance requirements of the job could be assessed through subsequent physical demands analyses and ultimately the determination of minimum occupational fitness requirements for UK firefighting roles. In order to ensure the safety of workers in physically demanding safety-related jobs, employers must have an understanding of the arduous nature of the roles undertaken by employees. This is determined by conducting a job, or task analysis which often involves collecting a combination of objective, evidence-based and subjective information (Constable & Palmer, 2000).

Previous task analysis studies have used a variety of established methods such as workplace observations (Bos et al., 2004) and survey response data from a sample of the workforce (Office of the Deputy Prime Minister, 2004a) to understand the nature of specific occupations. For this study a workplace observation study would not have been suitable due to the unknown timing of emergency incidents. As such, some of the most critical and/or physically demanding aspects of the role may not be captured by this type of analysis. Additionally, whilst survey data can involve large numbers, which are often representative of the workforce, we utilised open discussion and blinded voting with subject-matter experts, which aided navigation through previously identified potential sources of subjectivity within the task analysis process (Constable & Palmer, 2000; Tipton et al., 2012).

In this particular study, we identified a range of physically demanding tasks considered critical to incumbents in a firefighting role, which were casualty evacuation; equipment carrying; hose running; stair climbing; wild-land firefighting and the lifting, extending and lowering fire service ladders. These activities are similar to those reported previously in the UK fire and rescue service (Blacker et al., 2015; Rayson et al., 2009a) and are comparable to tasks performed by other firefighting populations (Bos et al., 2004; Elsnor, 2008; Holmer & Gavhed, 2007; Von Heimburg et al., 2006). Tasks that involve, walking, running and climbing stairs combined with having to move heavy equipment and/or casualties whilst wearing restrictive personal protective equipment remain important components of the firefighter role, all of which interact to elicit a substantial physical demand upon incumbents. This consistency with other firefighting populations

and the experience of the subject-matter experts used in this study lend confidence that the resultant tasks are representative of the occupation. The analysis of occupational roles within this study has gone further than many other task analysis studies by identifying specialist roles and determining the critical and most arduous generic tasks of all the recognised occupational roles within the UK fire and rescue services.

Although adding female subject-matter experts to the panel would have been more favourable, utilising a panel of experienced personnel in this study facilitated the understanding of the current practices adopted in the UK fire and rescue service. This would be effective for determining the method of best practice for any physically demanding occupation. Practical knowledge of manual handling guidelines, standard operating procedures and health and safety regulations assisted in the development of realistic single-person simulations of occupational tasks that accurately reflected job requirements. This is vital for correctly assessing the physical demand of a task and, when developing subsequent simulations, maintaining external validity (Tipton et al., 2012). Examples from the current study included ensuring employee safety by not expecting any firefighter to run whilst wearing breathing apparatus equipment and adhering to a manual handling regulation maximum carrying weight of 25 kg (Health & Safety Executive, 2004). The TP were also instrumental in developing realistic scenarios for each of the tasks. As the successful completion of many firefighting activities are recognised as being time-sensitive, it was important that a detailed scenario for each task was identified in order to clarify the situational context/urgency of that task with a view to minimise potential subjectivity when identifying what was an acceptable or unacceptable speed of performance.

Extant research examining occupational physical demands has often required participants to perform tasks as quickly as possible (Holmer & Gavhed, 2007; Von Heimburg et al., 2006). Other researchers have investigated demands based on a pace self-selected by participants using their experiential judgement of an emergency situation (Jamnik, Thomas, Shaw, et al., 2010). Whilst it may be important to recruit current trained employees as participants in such studies, it may not be appropriate to assume that all incumbents have maintained role specific fitness levels to carry out these tasks at an acceptable pace. This is particularly relevant in the fire and rescue services where

physically demanding emergency calls are so infrequent that the job demands themselves appear to be insufficient for maintaining role specific fitness levels (Office of the Deputy Prime Minister, 2004b). Additionally, in many instances, the aims of the above task analyses have been solely to understand the physical nature of a job by observing employees in their uncontrolled work environment.

However, if a research project (such as a physical demands analysis) aims to quantify the physical fitness requirement associated with minimum acceptable job performance, very clear and distinct consideration should be given to controlling the pace at which incumbents perform job tasks to a minimum acceptable standard (Bilzon et al., 2001). If these considerations are met when completing an initial task analysis, any subsequent physical demands analyses can be conducted with consistent paces and performance standards. Controlling tasks to a constant predetermined pace also avoids a number of potentially confounding factors to eventual physical demand measurement such as participant physical fitness determining the physiological demand of the work performed (Tipton et al., 2012). For these reasons, the project team used video footage of each simulation being performed at set work rates allowing the subject-matter experts to review and clearly identify the minimum acceptable performance requirement for each activity in a fashion similar to the Bookmark method (Lewis, 1999). This would be an important consideration when developing minimum physical fitness standards for any physically demanding occupation where task performance is time-sensitive.

Whilst every attempt was made to develop a consultation process that dealt with subjective components of this analysis in a structured way, it is clear that when running focus groups with experienced subject-matter experts, some differences of opinion on the nature of the occupation and which tasks are most arduous may still arise especially if it had involved female panel members. Theoretically, these could be founded on differences in the particular occupational environment or geographical location in which the panel member works; their number of years of experience or their interpretation of the particular scenario(s) presented, including sex and age-related considerations. For instance, the minimum acceptable pace for the wild-land fire task received a polarised vote which could indicate a need to re-consider the appropriateness or design of the task or removal from the analysis altogether. As such, one of the limitations of this study, which we would

seek to address in any future studies, was that no female personnel volunteered to participate in the technical panel or the filming of task simulations and that clarification should be sought on the inclusion/exclusion of any tasks that vary widely in employee practice.

Utilising a group of industry stakeholders to subsequently endorse the decisions made throughout the project may have increased the ecological validity of the outcomes from open discussion. However, analysis of reliability of the task- and pace- selection process were not conducted. As such, the research could be further improved with the inclusion of a test-retest of the voting process, and subject matter experts retrospectively endorsing trained incumbents at the selected paces to be “safe and efficient”. Finally, it should be acknowledged that other activities such as using heavy equipment at road traffic collisions or water rescue activities were also identified as physically arduous tasks for UK firefighters but were not included on the basis that they are sometimes specialist, as opposed to generic, tasks. However, these emergency incidents are not uncommon and, due to their importance, it would be favourable for firefighters to be physically capable of working at such incidents and may therefore warrant further investigation.

This study completed a rigorous task analysis of the critical and most arduous activities undertaken by UK fire service personnel, using a logical, systematic and structured format and engaging subject-matter expertise from within the organisation. This, in conjunction with a blinded voting format and constructed videos of firefighting activities, allowed for the effective determination of the minimum acceptable performance standards. Including a more divergent subject-matter expert panel with respect to age, sex and race, the structured steps identified within this task analysis methodology could be employed to establish minimum physical employment standards for other physically demanding public safety occupations.

CHAPTER 4

Study 2

Development of Role-related Minimum Cardiorespiratory Fitness Standards for Firefighters and Commanders

4.1 Introduction

The role of a firefighter requires a relatively high level of cardiorespiratory fitness to perform operational tasks safely and effectively (Gledhill & Jamnik, 1992a; Von Heimburg et al., 2006). Cardiac events during emergency incidents account for the largest number (45%) of on-duty firefighter fatalities in the United States (Kales et al. 2007; Fahy, LeBlanc, and Molis 2013), and so firefighters with poor physical fitness and cardiovascular health may be at increased personal risk when performing occupational duties. However, physical fitness standards are not consistently based on metabolic demand of essential occupational tasks.

Direct measurement of oxygen uptake in firefighter populations indicate the metabolic demand is rarely below $35 \text{ mL}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$ (Sothmann et al., 1990), and consistently in excess of $40 \text{ mL}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$ (Bilzon et al., 2001; Gledhill & Jamnik, 1992a; Von Heimburg et al., 2006). Physical demands analyses in firefighters, however, are typically limited to estimating (but not directly measuring) metabolic demand, indicating cardiovascular strain of between 60-95% of maximum (Eglin, Coles, & Tipton, 2004; Richmond, Rayson, Wilkinson, Carter, & Blacker, 2008). In addition, the majority of work has utilised tasks paced by the participant, where measured physical demand is altered by individual effort and absolute work capacity (Lemon & Hermiston, 1977). Notably, in one study using Naval firefighters metabolic demand was measured using tasks with constant predetermined “acceptable” paces and protocols designed to elicit a representative peak demand (Bilzon et al., 2001). Reproduction of this study design, in combination with further considerations from the gold-standard process for developing role-related physiological employment standards (Jamnik et al., 2013; Tipton et al., 2012), should therefore be replicated in civilian (non-military) firefighters. In addition, the physical roles expected of different ranks within those working at emergency incidents may have disparate physical responsibilities and therefore different physical demand. This would suggest different fitness standards would be employed for generic

‘operational’ firefighters tackling the incident and for those in an ‘incident command’ role at the scene. However, the predominance of physical demand research has not attempted to assess these differences, electing to solely examine the demands upon the firefighter.

The primary aim of this study was to quantify the peak oxygen cost (and therefore metabolic demand) of several simulated firefighting tasks, performed to a minimum acceptable and representative standard for firefighters and, separately, for those in incident command. The secondary aim was to derive minimum cardiorespiratory fitness standards for safe and efficient work in these two roles using these data, in conjunction with well-established methods for the development of physiological employment standards.

4.2 Methods

Best practice for the development of fair and justifiable occupational capability tests and minimum physical fitness requirements was recently reviewed by Tipton, Milligan & Reilly (2012) and the recommendations implemented for this study. A technical panel of 13 highly experienced fire service personnel (comprising one firefighter, two crew managers, five watch managers and five station managers from 10 UK fire and rescue services with an average of 17 years of experience) was established to provide information on the composition and best practice of firefighting tasks. This panel was assembled by requesting nominations for highly experienced incident managers from the UK fire and rescue services. Though a gender- and sex- diverse panel would have been preferable, unfortunately no female personnel came forward to sit on the panel.

Through a series of focus groups, the panel was consulted to identify the most physically arduous generic tasks that all firefighters must be able to perform. In the case of UK firefighters, this includes both the duties of structural and wild-land fire suppression. Subsequently, with further consultation, single-person simulations of five tasks were designed using the criteria that each simulation should: (a) replicate and employ best practice for the criterion tasks; (b) reflect one individual’s responsibility

within the task; (c) be reproducible and standardised in nature; (d) be long enough to elicit a steady state of metabolism during exercise and; (e) be completed at a “minimum acceptable pace” agreed by the technical panel in order to assess minimum occupational demand and standardise task intensity. These focus groups included open discussion and agreement of the typical distances covered during these specific tasks during an operational incident, whilst adhering to health and safety regulations. The agreed distances subsequently became those employed in the single-person simulations designed.

To select the minimum acceptable pace, a voting system was used in which panel members were shown a set of videos of each task at a variety of calculated speeds, and were asked to vote anonymously for the minimum expectation for safe and efficient task completion. The speeds were determined using the bookmark method (Rogers et al., 2014), specifically by examining the typical speed of a training instructor performing the tasks, then having the task completed one integer of speed faster and slower than this pace (giving a “slow”, “moderate” and “fast” pace). Voters were also given an option of choosing a speed half way between these, giving five choices in total. Before each video, a contextual scenario was given to control for the perception of the intensity and/or urgency of each task. The panel also identified which tasks would also be completed by ‘incident commanders’ and if/how these would differ from the physical responsibilities of generic firefighters. If differences were identified, separate addition single-person simulations of tasks were then designed to be appropriate for the ‘incident commander’ role.

Participants

Sixty two (50 male, 12 female) operational firefighters (Table 4.1) attended the Fire Service College (Moreton-in-Marsh, Gloucestershire, UK) and gave written informed consent to take part in the study following a full written and verbal brief. Inclusion criteria were that participants were trained, currently operational and deemed medically fit for service. Estimated maximum oxygen uptake data from ramped submaximal treadmill test (Buckley et al. 2004) were extracted from personnel records but were only available for 60 of the 62 participants (mean \pm SD: $50.0 \pm 6.6 \text{ ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$

¹). The study was approved by the University of Bath's Department for Health Research Ethics Committee (Reference number: EP 12/13 6).

Table 4.1. Participant Characteristics. Data are mean (\pm SD).

Variable	Participants		
	All ($n=62$)	Male ($n=50$)	Female ($n=12$)
Age (yr.)	40 (± 10)	40 (± 10)	37 (± 7)
Body mass (kg)	80.8 (± 11.8)	84.1 (± 10.1)	67.0 (± 8.0)
Height (m)	1.76 (± 0.07)	1.78 (± 0.06)	1.69 (± 0.05)
BMI ($\text{kg}\cdot\text{m}^{-2}$)	26.1 (± 3.2)	26.7 (± 3.0)	23.6 (± 2.7)
Estimated body fat (%)	21.8 (± 5.6)	20.6 (± 4.7)	26.9 (± 6.4)
Estimated VO_2 max ($\text{ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$)	50.0 (± 6.6)*	50.9 (± 6.3)	46.6 (± 7.4)

*Estimated maximum oxygen uptake data ($n=60$), not available for two male participants.

Study protocol

Anthropometric data (body mass, height, estimated body fat (BodyStat 1500, Bodystat Ltd., UK)) were obtained in the morning prior to the physical demands analysis. Participants were then randomly assigned to complete five simulated firefighting tasks with at least 60 min of recovery separating each task. Before each task a full verbal brief of the simulation was given. During each task a project researcher moved with the participant and gave instruction. Throughout the trial day participants were allowed access to food and drink *ad libitum*.

Four out of the five tasks were paced by audible beeps that corresponded with cones placed at five metre intervals. The stair climb was paced by a metronome where each sound corresponded to one step, played to the participant via headphones. All tasks were completed in full personal protective clothing consisting of helmet, shirt, tunic, leggings, boots, gloves (Mass of ensemble: 8.2 kg), with the exception of the wild-land fire task where tunic and helmet were not worn. In two tasks, the stair climb and casualty evacuation, a rucksack was also worn equivalent to the mass of a self-contained breathing apparatus unit (12 kg). The various equipment worn for each simulation was in keeping

with what would realistically be worn during an emergency incident according to the subject-matter expert technical panel.

Task descriptions

For the purpose of the task descriptions a “length” is a traversal of the 25 m course. All tasks were conducted in accordance with UK manual handling regulations, which stipulates maximum manual lifting of 25 kg. Note that both the stair climb and wild-land fire tasks each contain two sections that simulate, separately, the roles of a generic firefighter and an ‘incident commander’. From the minimum acceptable pace and distance, a minimum expected time for task completion is also stated below.

The hose run task was designed to simulate an operational scenario of establishing a water supply between a fire engine and a fire hydrant 100 m apart by carrying and ‘running out’ four 25 m standard issue 70 mm hoses (13 kg each). ‘Running out’ hose consists of the firefighter jogging while holding the hose in front of them at chest height on an attached spindle – letting the hose unravel as they traverse the course. The task was completed on a straight 25 m course at 8 km·h⁻¹ over a total distance of 700 m. This adhered to the realistic scenario of advancing to, and returning from, the hydrant (200 m) at both the start and end of the task, in combination with the most efficient procedure for ‘running out’ four hoses by the conclusion of the task. The task consisted of (in this order): 8 x unladen lengths, 3 x lengths carrying two hoses, 1 x length carrying one hose, 2 x lengths running out two hoses, 2 x unladen lengths, 1 x length carrying two hoses, 1 x length carrying one hose, 2 x lengths running out two hoses, 8 x unladen lengths. Minimum expected task duration: 5:05 min.

The equipment carry task was designed to simulate the single-person contribution to a team manually transporting a portable pump (or similar) over 200 m and was completed on a straight 25 m course at 5.5 km·h⁻¹. The participant carried a 25 kg barbell 8 x lengths, and was allowed to place the weight down to shift grip if necessary as long as pace was then recommenced in a timely manner. Minimum expected task duration: 2:11 min.

The stair climb task was designed to simulate a single-person contribution to a breathing apparatus team (i.e.: wearing a breathing apparatus unit; 12 kg) carrying a high-rise pack (50 kg between two people) to an incident six floors above ground level. This task was completed at 95 steps·min⁻¹ in a high-rise stairwell which consisted of six floors, with two flights of stairs between each floor and 10 steps per flight where they fully ascended and descended the stairwell unladen then fully ascended the stairwell while carrying a dumbbell (25 kg), and descended again, unladen. Total minimum expected task duration: 6:04 min. The first half of the task was used to simulate the ‘incident commander’ role.

The casualty evacuation task was designed to simulate entry (with breathing apparatus unit; 12 kg) to a commercial property fire (Phase 1) and casualty evacuation (Phase 2) at speeds of 6 km·h⁻¹ and 3 km·h⁻¹, respectively. The task was completed around a 25 m square course, with a fire engine with charged hose reel (37 kg) and sledgehammer (4 kg) at one corner, and a dummy at the opposite corner (55 kg, which represents half of the 90th percentile of the body mass of the UK population) (Blacker et al., 2015). Total minimum expected task duration: 2:30 min. The participant started at the fire engine.

- 1) The participant (in this order) completed: 1 x length with sledgehammer, 1 x unladen length, 1 x length dragging charged hose, 1 x unladen length, 2 x lengths dragging charged hose (to approach the dummy). In this phase, the two sides of the square utilised were marked with cones every 5 m.
- 2) The participant (using standard procedure grip under armpits) dragged the dummy 2 x lengths (the final two sides of the square). In this phase, the lengths utilised were marked by cones at every 2.5 m (to elicit half the speed as the first phase, for the same bleeps).

The wild-land fire task was designed to simulate an individual’s contribution to a team fighting a wild-land fire. In the UK, this would consist of a fire on dry grassy terrain (not woodland) that is beaten using a standard issue fire beater (consisting of a long pole with an attached foam/rubber pad; total mass 5 kg) as part of a slowly advancing team of firefighters. This task covered a 400 m distance and was completed on a 50 m stretch of uphill undulating grassy terrain at 3.5 km·h⁻¹. The participant completed 2 x ascent and 2

x descent (200 m) of the course without a fire beater then completed 2 x ascent and 2 x descent again (200 m) while equipped with a fire beater. The fire beater was used to strike the ground on every alternate walking step during the final two ascents. Total minimum expected task duration: 6:52 min. The first half of the task (without the fire beater) was used to simulate the 'incident commander' role.

Task validity and authenticity

A series of questions were posed to participants at the end of each exercise to assess validity and authenticity of the tasks. Participants were asked a) whether they received adequate instruction, b) whether the task was an adequate reflection of what one might be expected to perform in a training or operational setting (*validity*) and c) whether, in their experience, the task pace was "too slow", "too fast" or "about right" (*authenticity*).

Measurement of physical demand

During each task, cardiovascular strain was measured at 5-s intervals by chest-mounted heart rate monitor (Polar, Finland) and oxygen uptake (VO_2) was measured continuously using portable breath-by-breath gas analyser (K4b², Cosmed, Rome, Italy). The K4b² was calibrated by performing room air calibration, reference gas calibration (O_2 15%, CO_2 5%) followed by the turbine calibration with syringe prior to capturing expired air measurements. Rating of perceived exertion was taken at the end of exercise using the Borg scale (Borg, 1982). To determine aerobic demand of the tasks, a minute of peak steady state VO_2 was selected for each participant within each task. Peak steady state was defined as the minute of oxygen uptake within the final two minutes of exercise which exhibited the fewest perturbations and which also did not appear to contain any substantial fluctuation in oxygen uptake. Each minute of steady state was cleaned for anomalous breaths by removing values above or below three standard deviations of the mean from that minute, and averaged for each task. For each steady state minute, average heart rate was also calculated. Resting heart rate was taken as the lowest heart rate observed during the entire day of data collection. Heart rate reserve (HRR) was then calculated by subtracting resting heart rate from age-predicted heart rate max ($220 - \text{age}$).

For each task, the steady state heart rate was also expressed as a percentage of heart rate reserve (%HRR).

Deriving a minimum cardiorespiratory fitness standard

Two different empirically-informed methods of deriving physiological employment standards were applied to the metabolic demand data. This was both to allow a means of comparison between methods, and in order to make an informed empirically valid decision on the appropriateness of the resultant standard. The first method subtracts one standard deviation of group metabolic demand from the demand of the most arduous measured task, with the rationale that 83.3% of the work force is then incorporated in the calculation of an employment standard (Jamnik, Thomas, Burr, et al., 2010). The second method uses the mean of the metabolic demands of the representative tasks, with the rationale that a) this would be closest to the cardiorespiratory level expected of an average participant without prior experience (assuming normally distributed data) and b) this mimics a generic emergency response by incorporating multiple occupationally-representative tasks (Bilzon et al., 2002).

To calculate the fitness requirement for each of these methods, a generic work-time relationship was used to estimate the work intensity that could be sustained for the duration of the given task(s) (Blondel, Berthoin, Billat, & Lensel, 2001; Louhevaara et al., 1986). For instance, if an individual should be capable of completing a 15-min task at approximately 90% of maximal oxygen uptake, the physical demand of that task represents 90% of the resultant minimum fitness requirement required for that task.

Data analysis

All statistical analyses were completed using IBM SPSS version 20 (IBM, New York, USA). Group averages were calculated for all variables. A one-way paired analysis of variance (ANOVA) with post-hoc Bonferroni adjustment was used to analyse differences, and locate variance, between physical demand characteristics of tasks. Sex was included as a between-subjects factor to assess any difference in physiological responses to tasks between males and females. When deriving the fitness standard, participants who did not complete a task, or did not keep to the designated pace for that

tasks were removed (only) from the data for that specific task. When comparing between tasks, ANOVA solely analysed those that completed every task successfully (n=47). Statistical significance was set at $p \leq 0.05$. All data are presented as mean \pm SD unless otherwise stated.

4.3 Results

Task validity and authenticity

All participants (100%) stated they received adequate instruction for each of the tasks. Almost all respondents (94%) stated that tasks were a valid reflection of what they might be expected to perform in training or operationally (Table 4.2). With the exception of the wild-land fire task, an average of 91% of respondents confirmed authenticity by agreeing the task paces were “about right”. The wild-land fire task was the only task not to be perceived as valid by more than 90% of respondents (84%), and to have work rate deemed too slow to be authentic by the majority of participants (52%).

Table 4.2. Participant responses to the instruction received, authenticity and pace of each task as minimum acceptable performance.

Task	Instruction (%)	Validity (%)	Authenticity of task pace (%)		
			“Too fast”	“About right”	“Too slow”
Hose Run	100	90	9.8	90.2	0
Equipment Carry	100	97	0	88.7	11.3
Stair Climb	100	98	1.6	96.8	1.6
Casualty Evacuation	100	98	1.6	88.7	9.7
Wild-land Fire	100	84	0	(32.3)*	(51.6)*

**Note: For wild-land fire task, percentages do not include the proportion of respondents (16.3%) who felt that the allotted pace of 3.5 kmh⁻¹ was suitable for during the beating activity but was inappropriately slow during regular unladen walking.*

Firefighter task performance

In the hose run 52 of 62 participants completed the task correctly (83.9%), with nine individuals completing the task but at an incorrect pace (six too slow and three too fast), and one unable to complete (ankle soreness). Three individuals in the equipment carry, and four individuals in each of the stair climb and casualty evacuation tasks did not complete the tasks correctly and/or erratically changed work rate. The wild-land fire task was completed successfully by all participants. In total, 47 individuals completed every task at the correct pace successfully and could be included in the statistical comparison between tasks.

Incident commander task performance

While all participants completed the ‘incident commander’ portions of the stair climb and wild-land fire tasks, one individual did not have sufficient data to analyse physical demand during the stair climb and was removed from the “incident commander” physical demands analysis.

Firefighter physical demand

Examining each task separately (by including all successful completers for each individual task), physical demand was measured by mean \pm SD peak steady state oxygen uptake ($\text{ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$) and percentage of estimated VO_2 max for hose run (47 ± 7 ; $94 \pm 15\%$; $n=52$), equipment carry (29 ± 4 ; $58 \pm 11\%$; $n=59$), stair climb (41 ± 7 ; $83 \pm 15\%$; $n=58$), casualty evacuation (36 ± 7 ; $72 \pm 13\%$; $n=58$) and wild-land fire (29 ± 5 ; $59 \pm 13\%$; $n=62$). Statistical comparison between tasks ($n=47$) revealed that the hose run task elicited significantly higher mean peak metabolic demand than all other tasks ($p<0.01$), whilst wild-land fire and equipment carry tasks both elicited the lowest relative to the other three tasks ($p<0.01$; Table 4.3). Mean \pm SD heart rate responses were different between tasks ($p<0.01$), with the hose run eliciting the highest cardiovascular strain ($171 \pm 11 \text{ beats}\cdot\text{min}^{-1}$) and wild-land fire the lowest ($137 \pm 14 \text{ beats}\cdot\text{min}^{-1}$). Similarly, the hose run and stair climb elicited the highest percentage of heart rate reserve, with $92 \pm 7 \%$ and $88 \pm 10 \%$, respectively, and wild-land fire lowest ($64 \pm 10 \%$). Perception of exertion differed between each task ($p<0.05$), increasing in corresponding order to measured

physical demand (Table 4.3). Metabolic demand did not significantly differ between male and female firefighters within any task ($p>0.05$; Table 4.3).

Table 4.3. Metabolic demand, cardiovascular strain and perceived exertion for peak steady state during firefighting tasks for participants who completed all tasks successfully (n=47).

Task	Minimum expected task duration (min:s)	Physical demand				
		VO ₂ (ml·kg ⁻¹ ·min ⁻¹) Mean (±SD)	% estimated VO ₂ max(±SD)	Heart rate (beats·min ⁻¹) Mean (±SD)	%HRR (±SD)	RPE (/20) Mean (±SD)
Hose Run	5:05	47 (±8)*	93 (±15)*	171 (±11)*	92 (±7)*	15 (±2)
Equipment Carry	2:11	29 (±5)	57 (±11)	141 (±16)*	68 (±13)*	11 (±2)
Stair Climb	6:04	42 (±7)*	82 (±15)*	166 (±13)*	88 (±10)*	14 (±2)
Casualty Evacuation	2:30	36 (±6)*	70 (±12)*	159 (±13)*	82 (±9)*	13 (±2)
Wild-land Fire	6:52	29 (±5)	57 (±12)	137 (±14)*	64 (±10)*	9 (±2)

*Table shows oxygen uptake (VO₂), heart rate, percentage of heart rate reserve (%HRR) and rating of perceived exertion (RPE). Minimum expected task duration shown is calculated from distances and the assigned constant speeds of the task simulations. *denotes that mean values are significantly different from all other tasks ($p < 0.05$) by two-way mixed model ANOVA (n=47).*

Table 4.4. Peak steady state metabolic demand data from those completing each task successfully, organised by sex.

Task	<i>n</i> (Male/Female)	VO ₂ (ml·kg ⁻¹ ·min ⁻¹) Mean (±SD)	
		Male	Female
Hose Run	52 (43/9)	48 (±7)	44 (±7)
Equipment Carry	59 (47/12)	28 (±4)	31 (±6)
Stair Climb	58 (47/11)	41 (±7)	40 (±7)
Casualty Evacuation	58 (48/10)	36 (±7)	36 (±6)
Wild-land Fire	62 (50/12)	29 (±5)	29 (±4)

Table 4.5. Metabolic demand, cardiovascular strain for peak steady state during simulated ‘incident commander’ duty within firefighting tasks.

Task	Physical demand			
	VO ₂ (ml·kg ⁻¹ ·min ⁻¹) Mean (±SD)	% estimated VO ₂ max	Heart rate (beats·min ⁻¹) Mean (±SD)	%HRR
Stair Climb	35 (±5) ^b	71 (±12) ^b	149 (±13) ^b	74 (±11) ^b
Wild-land Fire	23 (±3) ^{ab}	47 (±10) ^{ab}	124 (±15) ^{ab}	53 (±11) ^{ab}

Table shows oxygen uptake (VO₂), heart rate, percentage of heart rate reserve (%HRR). ^an=47, significantly different from all other tasks (p<0.05). ^bn=61, significantly different from the corresponding operational firefighter task (p<0.05).

Incident commander physical demand

When including participants that successfully completed the tasks, mean \pm SD oxygen uptake values for tasks that simulated an incident commander role were 35 ± 5 $\text{ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$ for the stair climb and 23 ± 3 $\text{ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$ for wild-land fire task. The physical demand characteristics used in statistical analyses for the ‘incident commander’ tasks are presented in Table 4.5. The physical demand of the stair climb and wild-land fire tasks simulating ‘incident commander’ duties were significantly lower than the two same corresponding tasks simulating the roles of operational firefighters ($p < 0.05$). The ‘incident commander’ simulation of wild-fire elicited significantly lower physical demand than all other simulations ($p < 0.05$). While the stair climb for ‘incident commanders’ elicited lower physical demand than the operational firefighter stair climb, the physical demand was statistically similar to the casualty evacuation task ($p > 0.05$).

Deriving cardiorespiratory fitness standards

One standard deviation below the metabolic demand of the most arduous task (hose run; $46.9 \text{ ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$) was $39.8 \text{ ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$. An exercise intensity of 95% VO_2 max was deemed sustainable for the duration of the task (Louhevaara et al., 1986), meaning that following the Jamnik et al. (2010) approach, the minimum cardiorespiratory fitness standard was calculated as $41.9 \text{ ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$ ($39.8 \text{ ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1} / 0.95$). Since authenticity and physical demand data suggested the wild-land fire task was not sufficiently representative of the occupational task, and thereby not externally valid, it was not included in the production of a fitness standard in the method described by Bilzon et al. (2002). Therefore, the mean of the four valid tasks (hose run, equipment carry, stair climb, casualty evacuation) was taken ($38.1 \text{ ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$). The summation of the minimum expected durations of these tasks (Table 3) was 15:50 min. This was deemed sustainable at an intensity of 90% VO_2 max (Louhevaara et al., 1986), therefore producing a resultant cardiorespiratory fitness standard of $42.3 \text{ ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$ ($38.1 \text{ ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1} / 0.90$). For individuals in ‘incident command’ roles, from utilisation of the stair climb alone (sustainable at 95% VO_2 max), minimum cardiorespiratory standards from the two methods were 31.6 and $36.8 \text{ ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$, respectively.

4.4 Discussion

Firefighters must possess adequate levels of cardiorespiratory fitness to meet the physical demands of arduous occupational tasks. Yet minimum occupational fitness standards are rarely directly derived from these demands to ensure empirical validity. The primary purpose of this study was to quantify the metabolic demand during minimum acceptable performance of representative occupational tasks of operational firefighters and, separately, “incident commanders”, to inform the implementation of minimum cardiorespiratory fitness standards. Group mean metabolic demand for firefighters ranged from 27 ml·kg⁻¹·min⁻¹ (wild-land fire task) to 47 ml·kg⁻¹·min⁻¹ (hose run), eliciting an average of 57% and 93%, of estimated VO₂ max and 64% and 92% of heart rate reserve. All tasks were agreed to be accurate representations of occupational duties by over 90% of study participants, with the exception of the wild-land fire task (84%). Using published methods for the development of physiological employment standards, a minimum cardiorespiratory fitness standard of 42.3 ml·kg⁻¹·min⁻¹ was calculated for safe and efficient work in operational firefighters, and 36.8 ml·kg⁻¹·min⁻¹ for ‘incident commanders’.

This investigation is the first to publish both occupation-specific metabolic demands for UK (non-military) firefighters and, by implementing pre-determined minimum acceptable paces during tasks, role-specific minimum cardiorespiratory fitness standards for this occupational group. The few previous physical demands analyses completed in UK firefighters published in peer-reviewed journals (Eglin et al., 2004; Richmond et al., 2008), have been limited to measuring cardiovascular strain, finding exertion of between 60-95% of maximum heart rate. By direct measurement of metabolic demand, Sothmann et al. (1990) observed an average metabolic demand of 30.5 ml·kg⁻¹·min⁻¹ during seven successive firefighting tasks in United States firefighters, representing 76% of the sample average VO₂ max (39.9 ml·kg⁻¹·min⁻¹). During stair climb and casualty evacuation tasks Gledhill & Jamnik (1992) and von Heimburg, Rasmussen & Medbø (2006) in Canadian and Norwegian firefighters, respectively, measured substantially higher metabolic demand values (44 ml·kg⁻¹·min⁻¹), similar to those in the present study. The above studies involved entirely self-paced tasks. Where minimum acceptable paces were employed, the average physical demand of a sample of UK

shipboard Naval firefighters over five tasks was 36.2 (range 23-43) $\text{ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$ representing between 44 and 82% of the average participant VO_2 max (Bilzon et al., 2001). Although these tasks were paced and designed to elicit a valid steady state of physical demand, it is evident that wide variation exists in the different occupational roles of firefighters, as well as the occupational requirements of different national and civilian/military services. As such, the specificity and experimental control implemented within the present study produce the first accurate articulation of the aerobic capacity required to safely and effectively complete the role of an operational UK domestic firefighter.

The tasks implemented in the current study compare favourably to those used in other firefighting physical demands analyses. Gledhill & Jamnik (1992) and Bilzon et al. (2001) investigated highest occupational applications of cardiorespiratory endurance and strength in Canadian and shipboard UK naval firefighters, respectively, incorporating combinations of hose manipulation; dragging a casualty; ladder and/or stair ascension and carrying heavy equipment over distance. In government research to develop point-of-entry standards for UK operational firefighters, a task analysis by Rayson et al. (2009) produced tasks including a shuttle-run based equipment carry and casualty evacuation scenario. These tasks, while specific to UK firefighters, were designed for entry-standard fitness testing, and not with the same specific aims of the current research (Blacker et al., 2015). Therefore, a task analysis specific to this study was required (Chapter 3). In addition to task selection itself, other design elements, such as the bookmark method (Rogers et al., 2014), were used to optimise relatedness to the job and accommodate the specific physical measurements needed to inform the development of a minimum cardiorespiratory standard (Jamnik et al., 2013; Tipton et al., 2012). Our study is the first to apply a voting system with subject matter experts, where voters were blinded to others' responses and to the actual speeds they were observing, to establish minimum acceptable paces for each physical task. The visual observation of a task at a variety of speeds without having the actual speed values available (i.e.: 6 vs. 8 $\text{km}\cdot\text{h}^{-1}$) facilitated the decision making process. This process (a) removes subjectivity of successful task performance, and (b) improves experimental control by establishing a steady state of oxygen cost, since

large alterations in pace or exertion during physical activity would likely introduce error in oxygen uptake.

Two empirically-informed methods for producing occupational fitness standards from physical demands analyses were applied to the data from the present study (Bilzon et al., 2002; Jamnik, Thomas, Burr, et al., 2010) with additional consideration of the relationship (of human limitations) between work-load and time. The two methods applied here produced remarkably similar results, lending confidence to the minimum bounds of cardiorespiratory fitness required for firefighting. The authors recommend that $42.3 \text{ ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$, derived from the mean of sustained metabolic demand of the representative tasks (hose run, equipment carry, stair climb, casualty evacuation), be the minimum cardiorespiratory capacity for operational UK firefighters. This is with the rationale that the occupational fitness standard is derived from a range of tasks, which incorporate different physical requirements of the occupation while also possessing high relatedness to the job and more closely mimicking a generic emergency response than the hose run task alone.

During emergency duties, it is typical for firefighters to move from one physical task to another with little to no break, akin to a circuit exercise scenario (Jamnik et al., 2013). In addition, the summed duration of the tasks (15:50 mins) is similar to mean duration observed for in-dwelling fire incidents during emergencies (14:20 minutes; Sothmann et al. 1990) and the practical limit of a standard breathing apparatus. As it is unreasonable for an individual to complete any task at maximal oxygen uptake for more than a few minutes (Blondel et al., 2001; Louhevaara et al., 1986), the resultant minimum fitness requirement of $42.3 \text{ ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$ surpasses the physical demand of the tasks and/or a generic emergency response scenario to the extent that the occupational duties can be completed safely for their expected duration (Billat & Koralsztein, 1996; Blondel et al., 2001; Gleser & Vogel, 1973). It is also expected, since humans can work supra-maximally (in relation to cardiorespiratory capacity) for short durations, that the period of peak steady state of $47 \text{ ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$ observed for approximately two minutes of the

hose run would still be achievable for firefighters who possess fitness at, or marginally above, the proposed fitness standard.

A novel aspect of this study was the examination of role-specific job requirements. The lower fitness standard of $36.8 \text{ ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$ we recommend for incident commanders reflects the identified differences in physical responsibility from those of generic firefighters. Physiological employment standards should be role-specific, based on successful performance of the minimum expected requirements of that occupation, irrespective of sex or any other individual characteristic. This is particularly relevant for an occupation where the physical requirements for men and women are the same, as is the case for UK firefighters. However, the comparison between males and females can be useful to highlight if female personnel complete tasks safely and efficiently at a significantly lower metabolic cost than male personnel (who would typically be larger, possess greater muscle mass and higher cardiorespiratory fitness). This could ultimately suggest that the minimum standard should be derived from the metabolic demand of single-sex groups.

However, in the present study there were no significant intra-sex differences in metabolic demand in any tasks. It was necessary to exclude any participant from the derivation of the fitness standards if prescribed task pace was not maintained, either by completing the task(s) too quickly ($n=5$), too slowly ($n=8$, 6 with available fitness data) or too erratically ($n=2$). With these small sample groups and only estimated aerobic capacity data it would be misleading to base assumptions on whether the fitness of these individuals was the cause for unsuccessful adherence to prescribed task speed. However, in light of our proposed VO_2 max standard for firefighters, it is interesting to note that the average estimated fitness levels of participants who performed tasks too quickly and too slowly were $54 \text{ ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$ and $40 \text{ ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$, respectively.

It is worth noting that the cardiorespiratory fitness data in this study were estimated, rather than directly measured, and data were not available for two participants, which may limit the interpretation of the relative physical demands data, as well as comparisons with data from other firefighter populations. The comparatively low validity

and authenticity scores for the wild-land fire task indicate the task was not of a suitable level of ecological validity, where the shortened task simulation was not able to replicate the fatigue experienced in the real-world example, which would likely endure for several hours over much longer distances. Such a task was therefore not considered suitable, nor practical, for a controlled testing environment. Similarly, a challenge for any such trial is that it is not possible to safely replicate other stressors in the operational environment (i.e. fire, heat, smoke, darkness).

However, it is evident that in healthy, euhydrated adults environmental heat stress and subsequent hyperthermia, has greater deleterious effect on altering perception of exertion (i.e. RPE) driven, in part, by protective cerebral mechanisms (Nielsen & Nybo, 2003; Nybo, 2007), than on actual metabolic rate. While absolute metabolic rate remains similar during exercise in the heat, the ability to attain VO_2 max may be attenuated (Arngrimsson, Stewart, Borrani, Skinner, & Cureton, 2003). Performing a task at a workload relative to VO_2 max in the heat can, therefore, cause metabolic demand to appear higher than when the same work is performed in a thermoneutral environment. In reality, only approximately half of one task in the present study (i.e. casualty evacuation) would be completed under thermal stress in a real world setting. This task would not be expected to be performed near participant VO_2 max (group mean metabolic demand: $36 \text{ ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$), nor expected to result in significant hyperthermia. As such, including a task in excessive heat would likely not have affected the overall metabolic demands observed, and instead have reduced the safety and reliability of the experimental conditions. It is also acknowledged that the varied role of a firefighter requires components of fitness that extend beyond cardiorespiratory fitness, and while the tasks included require a wide variety of applications of fitness, further work should take a similar approach to investigate the muscular strength and endurance requirements for specific operational firefighting tasks. Validation of such a battery of fitness tests would provide a more holistic understanding of the physical requirements for operational firefighting.

This study supports the rationale for conducting physical demands analyses to derive evidence-based occupational fitness standards for arduous occupations. The metabolic demands exhibited were similar to those observed in other firefighting

populations. We used well-established methods to determine that a cardiorespiratory fitness standard below $42.3 \text{ ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$ would not be commensurate with safe and efficient operational performance and that a lower standard of $\geq 36.8 \text{ ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$ would be appropriate for ‘incident commanders’. This study supports the implementation of routine assessments of minimum cardiorespiratory fitness standards for operational firefighters and their incident commanders.

CHAPTER 5

Study 3

Establishing the Validity and Reliability of a Firefighting Simulation Test

5.1 Introduction

Physical fitness tests in demanding occupations typically take the form of either criterion (i.e. job simulation) or surrogate (i.e. predictive) tests. Criterion tests are either discrete job tasks or simulations, whereas surrogate tests measure components of fitness that are associated with performance on job-related tasks (Milligan et al., 2016). In an occupational setting, the type of test used will often depend on practical factors such as resource availability, financial constraints and/or safety considerations. However, since the aim of the fitness test is primarily to determine potential operational performance and, ultimately, suitability for employment, it is essential that the psychometric properties of the tests demonstrate an acceptable level of validity and reliability.

Whilst criterion tests naturally have content or ‘face’ validity, performances on these tests can be more markedly affected by factors such as weather conditions (where job simulations are performed outside) and test familiarisation than predictive tests (Boyd, Rogers, Docherty, & Petersen, 2015). Surrogate or predictive fitness tests (e.g. the Harvard step test) tend to be less complex, and can often be more easily administered in controlled conditions (i.e. a fitness or occupational health facility), which typically increases safety and test-retest reliability (Buckley et al., 2004). However, predictive tests inevitably contain prediction error which, when applied to a workforce, can introduce bias and call into question their validity and occupational relevance when compared to criterion tests that more closely resemble the job. Identifying valid and reliable fitness tests for physically demanding jobs is important to help identify workers that can undertake their role effectively thus improving employee and public safety.

Several studies have investigated the association between a variety of fitness indices with firefighting performance. Expectedly, cardiorespiratory fitness exhibits strong associations with tasks of longer duration (i.e. > 30 seconds), where little to no correlation is observed with individual, short duration (< 30 seconds) firefighting tasks (Rhea et al., 2004; Williford et al., 1999). Weak to moderate (but significant) correlations

have been observed with combined task simulations lasting between 5 and 10 minutes in U.S ($r = -0.38$) (Williford et al., 1999) and Norwegian ($r = -0.53$) firefighters (Von Heimburg et al., 2006). Typically, cardiorespiratory fitness exhibits the strongest associations with firefighting simulations lasting approximately 10 minutes ($r^2 = 0.57$) (Williams-Bell et al., 2009) or longer ($r = -0.72$) (Von Heimburg, Medbo, Sandsund, & Reinertsen, 2013), supporting the notion that cardiorespiratory fitness is an important determinant of task performance during longer duration tasks.

Studies investigating the reliability of timed task simulations have, however, received comparatively little attention. One study investigated the variability in task performance during a Canadian Forces firefighter work simulation test (Boyd et al., 2015). Participants completed six best-effort attempts of a standardised firefighting simulation with 24-48 hours between trials. Even with thorough orientation procedures and controlled test conditions, a continual improvement in performance on the occupational task was observed. Whilst the test-retest correlations between sequential trials were high ($r = 0.957$ to $r = 0.988$), significant variations in mean task performance were evident between all trials (Boyd et al., 2015).

In a series of recent studies (Chapters 3-4), we identified and established job-related simulations involving the critical and most physically demanding firefighting tasks (Stevenson, Siddall, Turner, & Bilzon, 2016). We also quantified the metabolic cost of these simulations and proposed a minimum cardiorespiratory fitness standard for operational performance (Siddall, Stevenson, Turner, Stokes, & Bilzon, 2016). However, it is unknown whether a simulation composed of these tasks can be used as a valid and reliable criterion test of operational fitness. This study was performed to: (a) assess the validity of a firefighter simulation test (FFST) to estimate maximal cardiorespiratory fitness (VO_{2max}) and; (b) establish the test-retest reliability of the FFST.

5.2 Methods

Subjects

Sixty-nine (64 male, 5 female) operational firefighters from seven UK fire and

rescue services volunteered for this study and gave written consent to participate following written and verbal explanation of the test procedures. All study participants were trained operational personnel and considered medically fit for firefighting duties. The study was approved by the university of Bath's Research Ethics Approval Committee for Health (REACH Reference number: EP 12/13 6).

Study design

Previous studies have determined the critical and most physically demanding tasks undertaken by UK firefighters and identified standardised simulations for each of these tasks (Stevenson et al., 2016). From this work a criterion firefighting simulation test (FFST) was developed by combining three of these tasks. These were selected based on the critical and physically demanding nature of the tasks and the ability to be easily replicated on a standard fire service training ground. After a familiarisation procedure, participants were required to complete this firefighting simulation in the quickest possible time along with a maximal treadmill test to determine their maximal oxygen uptake ($\text{VO}_{2\text{max}}$).

Firefighting Simulation

The firefighting simulation consisted of 'equipment carry', 'casualty evacuation' and 'hose run' tasks (Stevenson et al., 2016) and are described in detail in table 5.1. The tasks were selected to be completed in the order described (equipment carry, casualty evacuation, hose run) such that physical demand (reported previously) (Siddall et al., 2016) was incremental (i.e. least demanding to most demanding). Prior to undertaking the firefighting simulation, each firefighter was given full instruction of the task protocol and completed 2-3 attempts in the two weeks prior to the start of the testing procedure. On the day of assessment, participants were instructed to avoid strenuous exercise and to eat and drink as normal. Participants completed the FFST in full firefighting ensemble (i.e. tunic, leggings, boots, flash-hood, helmet, gloves (total mass ~8.2 kg)), whilst carrying a self-contained breathing apparatus set (total mass 12.0 kg) during the casualty evacuation component of the simulation.

Table 5.1. Description of the consecutive components of the firefighting simulation

Component	Description
1. Equipment carry	A simulated equipment-handling task carrying firefighting equipment. Performed by walking a flat 25 m course while carrying a 25 kg barbell. Consisting of completing 8, 25 m shuttles (total distance 200 m), followed by:
2. Casualty evacuation	A simulated entry to, and rescue of an unconscious casualty from, an industrial building. Consisting of: Dragging a charged hose reel 25 m; walking back 25 m; dragging another section of the hose 50 m; dragging a 55 kg dummy 50 m (total distance 150 m), followed by:
3. Hose run	A simulated water relay task to establish a water supply from a fire hydrant to a fire appliance 100 m apart using a total of four standard 70 mm hoses (weighing 13 kg) completed over a flat 25 m course. Consisting of: running 8 x 25 m, carrying 2 hose 75 m and 1 hose 25 m, rolling out 1 hose 25 m, then another hose 25 m; running 50 m, carrying 2 hose 25 m and 1 hose a further 25 m; rolling out 1 hose 25 m, then another hose 25 m; running 8 x 25 m (total distance 700 m).

Participants were asked to complete the FFST in the fastest time possible whilst adhering to standard operating procedures, manual handling and safety regulations. The time taken to complete each of the three stages/tasks of the FSST was recorded, as well as perceived exertion (Borg, 1982) at the end of the FFST.

Test retest reliability

At least 7-days later, a sub-sample of 22 participants (20 male, 2 female) completed a second best-effort attempt of the FFST to examine test-retest reliability. Both best-effort attempts were performed at the same fire station and at approximately the same time of day using the same equipment and pre-test conditions.

Laboratory cardiorespiratory fitness test

Participants also performed a maximal running protocol (in standard gym kit: running shoes, shorts and t-shirt) on a motorised treadmill (Life Fitness, Cambridge, UK) to determine their maximal oxygen uptake (VO_2max). This was conducted at the participants' designated fire stations in on-site gym facilities. Prior to the exercise test, anthropometric data (i.e. body mass, stature, estimated body fat) were collected. Body mass was measured on a flat digital scales (Seca, Germany). Stature was measured using a wall-mounted stadiometer (Seca, Germany) with the participant in bare feet, with heels buttocks and shoulder touching the wall with the head held in the Frankfurt plane. Estimated body fat percentage (BodyStat 1500, Bodystat Ltd., UK) was measured with the participant in supine position with legs and arms spread out, placing 2 electrodes onto the right foot (behind the second toes and on the ankle between the medial and lateral malleoli) and 2 onto the right hand (behind the knuckle of the middle finger and on the wrist next to the ulnar head). For the exercise test a 2-3 minute self-selected warm up was completed prior to commencing the test protocol which consisted of completing 4 (or 5 if required), 3-minute stages of walking or running at a constant speed with a 3% increase in gradient at the end of each stage until volitional exhaustion. Participants wore a portable breath-by-breath gas analyser (K4-B2, Cosmed, Italy). Peak VO_2max was established by calculating the mean of the highest minute of continuous oxygen consumption.

Statistical analysis

Statistical analysis was undertaken using SPSS version 22 (IBM, New York, USA). Mean anthropometric and performance ($\text{VO}_2\text{ max}$ and task time) data were compared between male and female firefighters using independent t-tests. A Pearson's correlation coefficient was used to determine the relationship between the FFST time and $\text{VO}_2\text{ max}$. Shapiro-Wilk's tests were conducted to assess for normality of data distribution. Standard error of estimate (SEE) statistics were calculated to determine the size of the mean error from the estimation plot. Test-retest reliability data were examined using Pearson's correlation coefficients between FFST attempts and a paired t-test was used to identify differences between mean performance times. Statistical significance was

set at $p \leq 0.05$. The variability between attempts was assessed using coefficient of variation (CV).

5.3 Results

The mean (and standard deviation) firefighter physical and performance characteristics are presented in table 5.2.

Table 5.2. Study participant physical and performance characteristics (mean \pm SD).

	All (n=69)	Male (n=64)	Female (n=5)
Age (y)	40 \pm 8	41 \pm 8	33 \pm 4*
Mass (kg)	85.8 \pm 12.8	87.2 \pm 12.0	67.2 \pm 5.4*
Height (m)	1.78 \pm 0.06	1.78 \pm 0.06	1.75 \pm 0.06
Body Mass Index (kg·m ⁻²)	27.0 \pm 3.6	27.4 \pm 3.4	22.0 \pm 2.5*
Estimated body fat (%)	19.7 \pm 5.5	19.4 \pm 5.5	23.4 \pm 4.7
VO ₂ max (ml·kg ⁻¹ ·min ⁻¹)	47.8 \pm 9.0	48.0 \pm 9.2	45.5 \pm 4.2
Simulation task time (s)	608 \pm 90	600 \pm 77	706 \pm 57*

* Significant difference between male and female firefighters ($p < 0.05$).

The male firefighters in this study were 8-years older ($p = 0.04$), 20-kg heavier ($p < 0.01$) and had a body mass index 5.4 kg·m⁻² greater ($p = 0.01$) than their female colleagues (Table 2). Interestingly, there were no significant differences between males and females in height ($p = 0.19$) or estimated body fat percentage ($p = 0.12$), despite the fact that males were 3 cm taller and 4 percentage points leaner than the females. The mean time to complete the firefighting simulation for all firefighters was 608 (\pm 90) seconds, with male firefighters (600 \pm 77 seconds) completing the task significantly quicker ($p < 0.01$) than the female firefighters (706 \pm 57 seconds).

Figure 5.1 shows the relationship between FFST task completion time (in seconds) and measured VO₂ max (maximal treadmill test). The time to complete the FFST was highly inversely correlated with cardiorespiratory fitness ($r = -0.73$, $p = 0.01$). The SEE was equivalent to 55 seconds on the FFST.

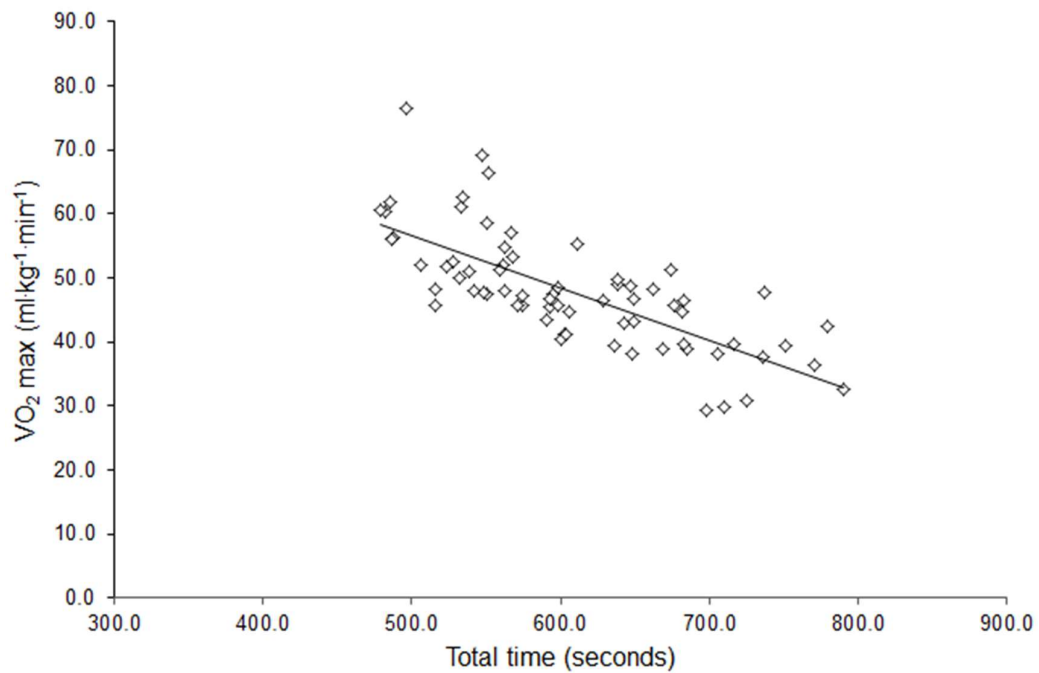


Figure 5.1. Relationship between completion time (seconds) on the firefighting simulation versus VO_2max ($n = 69$, $r = -0.734$, $r^2 = 0.539$), with line of best fit.

Data describing the physical characteristics of participants used in the reliability analysis are described in Table 5.3. The relationship between the FFST performance time trials are presented in Figure 5.2.

Table 5.3. Reliability study participants' physical and performance characteristics (mean \pm SD)

Variable	All ($n=22$)
Age (y)	42 ± 7
Mass (kg)	81.9 ± 11.6
Height (m)	1.78 ± 0.06
Body Mass Index (kg.m^{-2})	26.0 ± 3.0
Estimated body fat (%)	18.2 ± 4.9
Simulation task attempt 1 time (s)	612 ± 83
Simulation task attempt 2 time (s)	595 ± 74

Participants that performed the two reliability trials were, on average, quicker during the second trial ($595 (\pm 74)$ seconds) compared to the first ($612 (\pm 83)$ seconds), although this difference was not statistically significant ($p = 0.09$). Test-retest reliability of the FFST was high, revealing a strong relationship between the mean completion time of trial 1 and trial 2 ($r = 0.84, p = 0.01$). The coefficient of variation between the two tests was calculated to be 4.5%, equivalent to 27.1 seconds.

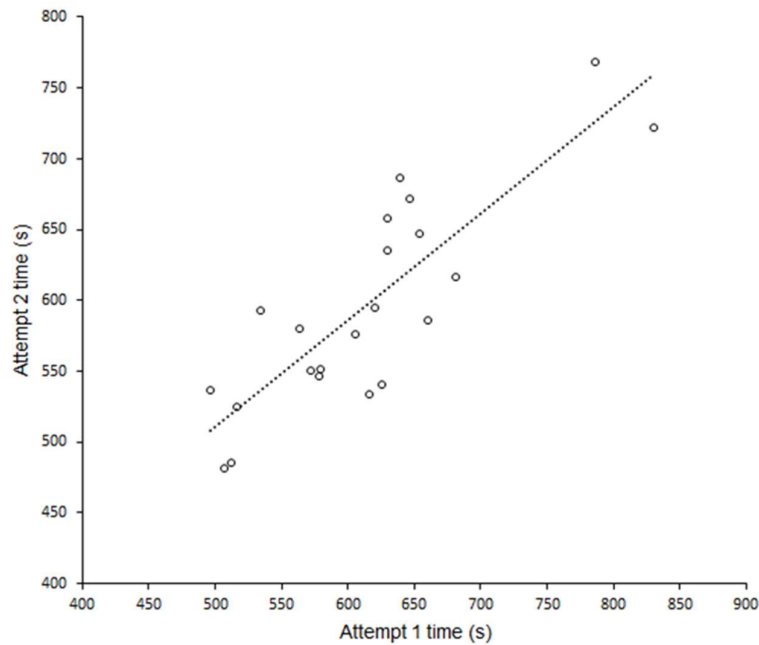


Figure 5.2. Test-retest reliability of the firefighter simulation test ($n = 22, r = - 0.842, r^2 = 0.708$).

5.4 Discussion

This study was undertaken to investigate the validity of a best-effort performance on a firefighting simulation test (FFST) to estimate the minimal maximal cardio-respiratory fitness (VO_{2max}) required for operational firefighting and to determine the test-retest reliability. The time taken to complete the FFST demonstrated a strong inverse correlation with VO_{2max} demonstrating that performance on the FFST is strongly determined by cardiorespiratory fitness. Furthermore, the repeatability of the simulation was highly reliable ($r = 0.842$). However, the error associated with the simulation (SEE of $6.13 \text{ ml.kg}^{-1}.\text{min}^{-1}$), and the test-retest coefficient of variation of 4.5% suggest it may

not be suitably accurate to supersede already widely-used predictive treadmill tests and may not be considered an appropriate criterion fitness test if used in isolation.

Whilst previous studies have identified equivocal findings in the relationship between firefighting task performance and cardiorespiratory fitness, test protocols used have varied widely (Von Heimburg et al., 2013). Studies investigating individual tasks or short-duration simulations have reporting little to no association with cardio-respiratory fitness (Rhea et al., 2004; Williford et al., 1999). However, work involving firefighting simulations lasting more than a few minutes have reported stronger and statistically significant associations with overall task performance (Von Heimburg et al., 2006; Williford et al., 1999), with the strongest correlations in firefighting simulations lasting greater than 10 minutes (Williams-Bell et al., 2009); (Von Heimburg et al., 2013). Interestingly, a number of these studies have also reported that when isolated from the complete simulation, specific tasks involving stair climbing (Williford et al., 1999) and casualty rescue (Von Heimburg et al., 2006) have elicited stronger correlations compared with the overall task performance. This supports the notion that the specific type or nature of the physical task(s) involved in the simulation may be as important as the duration of the task itself, in terms of determining the reliance on and strength of association with cardiorespiratory fitness. In support of this, stair climbing and casualty rescue tasks have been identified as being some of the most aerobically demanding tasks in firefighting (Elsner, 2008; Gledhill & Jamnik, 1992a; Siddall et al., 2016). These factors suggest that the combination of several different tasks, and subsequently, the total duration of the simulation in the current study may have facilitated the validity as an operationally relevant test of cardiorespiratory fitness.

Other methodological differences may also affect the relationship between the fitness variable and firefighting task performance. Two of the longest duration simulations described in the literature, used specifically designed firefighter applicant (Williams-Bell et al., 2009) or incumbent (Von Heimburg et al., 2013) simulation tests, where the former separated firefighting activities with periods of (walking) recovery (Williams-Bell et al., 2009). The nature of urgency by which the firefighting tasks are completed may also be of importance. Participants completing the simulations in the studies described were instructed to complete the tasks, “with no unnecessary waste of

time” (Von Heimburg et al., 2006), “at a steady and rapid pace” (Williford et al., 1999), or “as quickly as possible” (Rhea et al., 2004; Von Heimburg et al., 2013). It is therefore impossible to know whether all participants in these studies achieved best-effort performances, which would be more likely to elicit a stronger correlation with maximal cardiorespiratory fitness. In the present study, participants completed the FFST in the fastest possible time without walking or recovery between tasks. Furthermore, upon completion of the FFST, all participants reported an RPE of 17 or more. This represents the most likely reason why the relationship between task completion time and cardiorespiratory fitness in this study was stronger than those associations reported in the wider literature.

Despite the strong relationship between FFST performance and cardiorespiratory fitness observed in this study, only 54% of the total variance in task time could be explained by cardiorespiratory fitness. Additionally, the SEE of $6.13 \text{ ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$ equates to 55 seconds on the firefighting simulation. Whilst the SEE of 4.5% is similar to other reported studies (Williams-Bell et al., 2009), the value of this error in absolute terms indicates that this test may not be suitable to use as a stand-alone test to accurately determine fitness for duty in UK firefighters. The current minimum cardiorespiratory fitness standard for UK firefighters is $42.3 \text{ ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$, which would equate to a pass score of 10 minutes and 44 seconds on the FFST. However, the SEE equates to 55 seconds on the simulation test. As such, in isolation, this test may not be suitably accurate to determine whether a firefighter was fit or safe enough to successfully undertake firefighting duties without undue physical strain.

The test-retest reliability of the FFST was strong ($r = 0.84$) and the variability between the two trials was 4.5% or 27.1 seconds, which is similar to other studies involving firefighting simulations (Boyd et al., 2015; Von Heimburg et al., 2013). Whilst the correlations reported by Boyd et al. (2014) were higher ($r = 0.95 - 0.98$) between each consecutive pair of six trials and the variation between trials smaller (2.6%) during the latter part of the study, these tasks were completed 24 to 48 hours apart in an indoor, temperature-controlled facility. This was in contrast to the current study where trials were completed outdoors, one to two weeks apart and where natural variations in the ambient conditions were present, which likely explains the weaker correlations observed.

Considering the work of Boyd et al. (2015), improvements in best-effort times on the FFST may have been achieved where further task familiarisation was possible. However, as this study was conducted using operational firefighters, whilst on-duty, emergency duties and other work commitments made this difficult to achieve. Indeed, due to time restrictions, it was also not possible to collect strength or muscular endurance performance data from the firefighters used in this study. This may have been useful, allowing the determination of the effects of multiple fitness characteristics on FFST performance, particularly as both strength and muscular endurance have been reported to be important determinants of firefighting performance in specific tasks (Stevenson, Siddall, Turner, & Bilzon, 2017).

The FFST investigated in this study has been shown to be a reasonably valid and highly reliable test for predicting cardiorespiratory fitness. However, the error terms observed for the validity and reliability raise questions of the efficacy of the test when aiming to apply the minimum fitness standard for UK firefighters (Siddall et al., 2016). The use of this test in isolation and without prior health and fitness screening may not be suitable to effectively identify firefighters that are above/below the fitness standard and therefore capable of safe and efficient operational firefighting performance. Therefore, this type of test may be better utilised when included within a larger fitness management process to ensure firefighters are at least above a minimum level of cardiorespiratory fitness before undertaking this test. This would minimise the possibility for error and, more importantly, help ensure the safety of the incumbent performing the FFST. Importantly, the FFST can now be used as part of a regular operational firefighter fitness training programme.

CHAPTER 6

Study 4

Development of Minimum Muscular Strength and Endurance Requirements

6.1 Introduction

While many UK firefighters are subjected to routine fitness monitoring to ensure appropriate levels of fitness are maintained, fitness assessments almost exclusively focus on ensuring appropriate levels of cardiorespiratory fitness (Stevenson et al., 2009). Indeed, much of the international scientific research into minimum fitness standards for serving firefighters has tended to focus more on cardiorespiratory fitness requirements (Gledhill & Jamnik, 1992a; Scott, 1988; Sothmann, Landy, & Saupe, 1992) than on other components of physical fitness, such as muscular strength and endurance. This is despite a number of studies identifying their importance for safe and effective firefighting performance (Bilzon et al., 2002; Gledhill & Jamnik, 1992a; Jamnik et al., 2013). In the UK, minimum cardiorespiratory fitness standards for firefighters were recently identified and recommended (Siddall et al., 2016) but strength and muscular endurance standards for safe and effective performance of essential firefighting tasks remain unclear.

Studies comparing firefighting task performances with both laboratory (Lindberg et al., 2014) and gym-based tests of strength and muscular endurance (Henderson, Berry, & Matic, 2007; Michaelides et al., 2011; Rhea et al., 2004), have shown that physical ability tests can be used to predict firefighting performance. However, few investigations have both identified suitable surrogate tests (gym-based, easily replicable) and/or determined performance standards on these tests that are associated directly with minimal acceptable job performance. One study determined that a combination of three surrogate tests were able to predict performance on a fire suppression task (Sothmann et al., 2004). A combined test score was then validated against minimum acceptable performance standards previously identified through a job analysis process (Sothmann et al., 2004).

The authors identified that the derived cut score would correctly identify 89% of the ‘successful’ task performances and 72% of the ‘unsuccessful’ performances within the workforce. However, this study was conducted on firefighters from a single municipal

fire service in the USA and, to our knowledge, there are no other studies of this kind in other firefighter populations. In the UK, there is a lack of research investigating the minimum muscular strength and endurance requirements for performing the critical and most arduous firefighting tasks and/or using gym-based physical ability tests from which to derive physical employment standards.

The aim of this study was therefore to assess the sensitivity and specificity of common and replicable gym-based physical ability tests to predict performance of criterion operational firefighting tasks that require the largest application of physical strength and muscular endurance. To our knowledge, this will be the first study to identify and recommend minimum muscular strength and endurance tests and standards associated with minimal acceptable task performance for UK firefighting tasks.

6.2 Methods

This study was completed as part of a UK wide project to develop physical employment standards for incumbent UK firefighters. A task analysis process, using a best practice methodology (Tipton et al., 2012), identified the critical and most physically demanding generic tasks using muscular strength and endurance performed by all UK firefighters through consultation with a ‘technical panel’ of subject matter experts (Stevenson et al., 2016). Pilot testing was conducted at South Wales Fire & Rescue Service Training and Development Centre (Cardiff, UK) to determine the forces required to perform each of the identified tasks to a minimum acceptable standard (Stevenson et al., 2016), using standard fire service equipment. An analogue force dynamometer (Model 5002, Takei, Japan) was used to measure the force required to overcome inertia on each piece of fire service equipment involved in the individual tasks. For the ladder lift, the dynamometer was attached to the ladder and held over the ladder to determine the static weight / force required to lift the equipment. For the ladder pull activities, the dynamometer was attached to the ladder rope and pulled down to ascertain the static weight / force required to raise the ladder.

Following this, criterion tasks were either designed using these force measures or identified from previous research projects (Blacker et al., 2015) to simulate one individual firefighter's (single-person) requirement within each task. Best practice guidelines were adhered to in order that each task was performed safely and replicated the actual nature of the job (Her Majesty's Fire Service Inspectorate, 2004). To correspond with each criterion task, a gym-based physical ability test was identified. The criteria for selection of these tasks was that they used similar movements and/or application of force as their corresponding criterion tasks, used commonly available gym equipment and could be easily monitored (and safely controlled and/or 'spotted') by a practitioner. The criterion (occupational) tasks are described later, followed by their corresponding gym-based physical ability tests.

Participants

Twenty-six male (age 24 ± 5 y; mass 83 ± 15 kg; height 1.79 ± 0.07 m; BMI 26 ± 4 kg/m²; body fat 16 ± 5 %) and 25 female (age 24 ± 6 y; mass 63 ± 6 kg; height 1.65 ± 0.06 m; BMI 23 ± 3 kg/m²; body fat 26 ± 6 %) participants volunteered for this study and, after obtaining written and verbal explanation of the test procedures, provided informed consent to participate. Participants were recruited from two local universities as well as from support staff of South Wales Fire & Rescue Service. Since the tasks required no specialist skill or technique, operational firefighters were not recruited. The recruitment of civilians (non-firefighters) allowed similar proportions of male and female participants with divergent physical capabilities to be recruited. Participants completed a physical activity readiness questionnaire (Par-Q+) to ensure their safety to complete the physical tasks.

Protocol

Participants attended South Wales Fire & Rescue Service's Training and Development Centre, Cardiff, UK to complete the series of firefighting tasks and gym-based physical ability tests. Upon arrival, anthropometric measurements (body mass, height, estimated body fat (Bodystat 1500, Bodystat Ltd., UK)) were recorded for each participant. Following this, participants completed the occupational tasks and physical

ability assessments in a randomised order with adequate recovery between each task. All operational firefighting tasks were completed while wearing a standard firefighting ensemble (fire tunic, leggings, boots, helmet and gloves) to replicate the demands of working in firefighting equipment. The physical ability tests were performed in loose fitting gym clothing.

Criterion tasks

- *Ladder lift task* – The ladder lift task was completed using a bespoke fire service ladder lift simulator (Blacker et al., 2015). Participants performed the task by lifting a bar on a pivot arm from hip height to a height of 1.82 m, replicating half of the weight of the head of a 13.5 m fire service ladder (approximately 29 kg at the mid-lifting point). Participants completed a set routine corresponding to lift weights of 14 kg, 19 kg and finally 29 kg with two minutes rest between lifts. Task performance was recorded as a pass / fail to successfully lift the 29 kg to the required 1.82 m height in one complete motion
- *Ladder lower task* – The ladder lower task was completed using a wall-mounted PowerSport ladder simulator (PowerSport Fitness Ltd, Bridgend, UK). Participants were required to perform a single downward pull on the ladder rope with both hands from a vertically extended position to chest height in order to simulate the unhooking of the weight of a 13.5 m fire service ladder (approximately 42 kg) (Stevenson et al., 2016)
- *Ladder extension task* – The ladder extension task was completed using a wall-mounted PowerSport ladder simulator (PowerSport Fitness Ltd, Bridgend, UK). Participants were required to fully extend a 10.5 m fire service ladder at a set speed of 70 pulls per minute by continuously pulling (using a hand-over-hand action) on the ladder rope weighing approximately 28 kg (Stevenson et al., 2016)

Gym-based physical ability tests

All physical ability tests were preceded by a standardised warm up procedure (American College of Sports Medicine, 2010) and were separated by an adequate recovery period. Maximal performance on the physical ability tests that required a single transfer of force were assessed by one-repetition maximum (1RM) and for tests that

required repetitive motion, performance was assessed by number of continuous repetitions until volitional failure at a given load:

- *Seated shoulder press* – The seated shoulder press exercise (surrogate for the ladder lift task) was completed on a Body Solid power rack (Body Solid Ltd, Illinois, USA) using a standard Olympic bar with standard Olympic size weights in 2.5 kg increments. Participants were required to perform a 1RM overhead press whilst maintaining proper posture in an upright, seated position. The heaviest weight successfully pressed overhead was recorded
- *Seated rope pull-down (single)* – The seated rope pull-down exercise (surrogate for the ladder lower task) was completed on a commercial seated cable lat-pull down machine (Life Fitness Ltd, Illinois, USA). The lat-pull down bar was replaced by a section of standard fire service rope used for the extension of fire service ladders. Participants were required to perform a 1RM single pull down on the rope with both hands from a fully extended overhead position to chest level. The highest weight successfully pulled to chest height was recorded
- *Seated repeated rope pull-down (repeated)* – The repeated seated rope pull-down exercise (surrogate for the ladder extension task) was completed on a commercial seated lat-pull down machine (Life Fitness Ltd, Illinois, USA) using a set weight of 28 kg (corresponding to the weight of a 10.5 m fire service ladder). The lat-pull down bar was replaced by a section of standard fire service rope used for the extension of fire service ladders. Participants were required to repeatedly pull down on the rope with both hands to chest level and return to the starting position at a speed indicated by audible bleeps from a metronome, until failure. To correspond with the criterion ladder extension task, participants were instructed to time each downward pull and each return to starting position with a bleep set to 70 beats per minute (the minimum performance requirement identified by the technical panel), which equated to 35 downward pulls per minute. The test was stopped (and the number of repetitions recorded) when the participant was unable to complete a full repetition in time with the metronome or the participant could no longer maintain their grip on the rope

Statistical analyses

Independent t-tests were performed to identify the existence of significant differences in maximal performance in the physical ability tests between those who passed and failed the criterion tests and between males and females. Significance was identified as $p < 0.05$. For each criterion task the binary result (pass/fail) was plotted against the participants' maximal performance in the corresponding physical ability test. For each test, sensitivity (true positive rate) and specificity (false positive rate) were calculated at several hypothetical performance standards set at regular increments. Sensitivity, the ability of the predictive physical ability test to correctly identify those who passed the criterion test, was calculated using the following formula:

$$\text{Sensitivity} = \frac{TP}{TP + FN}$$

where TP denotes true positives, and FN denotes false negatives.

Specificity, the ability of the predictive physical ability test to correctly identify those who failed the criterion test, was calculated using the following formula:

$$\text{Specificity} = \frac{TN}{FP + TN}$$

where TN denotes true negatives and FP denotes false positives.

Accuracy was then determined by summing the number of true positives and true negatives and dividing by the total number in the population sample. Receiver-operating characteristic (ROC) curves were then plotted using the range of performance standards, with sensitivity on the y-axis and 1-specificity on the x-axis to determine the performance standard that was mathematically closest to maximising both specificity and sensitivity (perfect classification would be where both have a value of 1). Where applicable, this value was rounded to the nearest whole increment suitable for that performance measure.

6.3 Results

Thirty-one of the 51 participants (61%; 26 male, 5 female) successfully completed the ladder lift task. Thirty-nine (77%; 26 male, 13 female) successfully completed the criterion ladder lower task and 36 participants (71%; 25 male, 11 female) successfully completed the ladder extension task (Table 6.1). Significant differences in muscular strength were identified between the successful and unsuccessful groups in the ladder lift task (53 ± 13 kg vs. 25 ± 5 kg respectively; $p < 0.01$) and the ladder lower task (79 ± 20 kg vs. 48 ± 9 kg respectively; $p < 0.01$) and in muscular endurance on the ladder extension task (41 ± 22 repetitions vs. 13 ± 9 repetitions respectively; $p < 0.01$).

While male participants successfully completed all criterion tasks to the required standard apart from one individual who failed to complete the ladder extension task, a higher proportion of female participants failed to complete the ladder lift (80%), ladder lower (52%) and ladder extension (56%) tasks than those who were successful. The male participants in this study demonstrated significantly greater maximal strength compared to their female counterparts on the seated shoulder press exercise (55 ± 13 kg vs. 28 ± 8 kg; $p < 0.01$) and on the seated rope pull-down exercise (91 ± 14 kg vs. 52 ± 9 kg; $p < 0.01$) and greater muscular endurance compared to their female counterparts in the seated repeated rope pull-down exercise (49 ± 20 repetitions vs. 16 ± 9 repetitions respectively).

Table 6.1. Performance on the criterion tasks and the corresponding maximal performances (mean \pm SD) during the physical ability test(s).

Criterion task	Ladder Lift		Ladder Lower		Ladder Extension	
	Pass	Fail	Pass	Fail	Pass	Fail
n	31	20	39	12	36	15
Male	26	0	26	0	25	1
Female	5	20	12	13	11	14
Shoulder press 1RM (kg)	53 \pm 13	25 \pm 5*	-	-	-	-
Seated pull 1RM (kg)	-	-	79 \pm 20	48 \pm 9*	-	-
Repeated pull (reps)	-	-	-	-	41 \pm 22	13 \pm 9*

* - Significantly different from those that passed the criterion task.

Figures 6.1-6.3 (below) show individual performances (pass/fail) in the criterion tasks (ladder lift, ladder lower and ladder extension, respectively) versus maximal performances in the corresponding physical ability test (Panel A) and corresponding ROC curve derived from these data (Panel B).

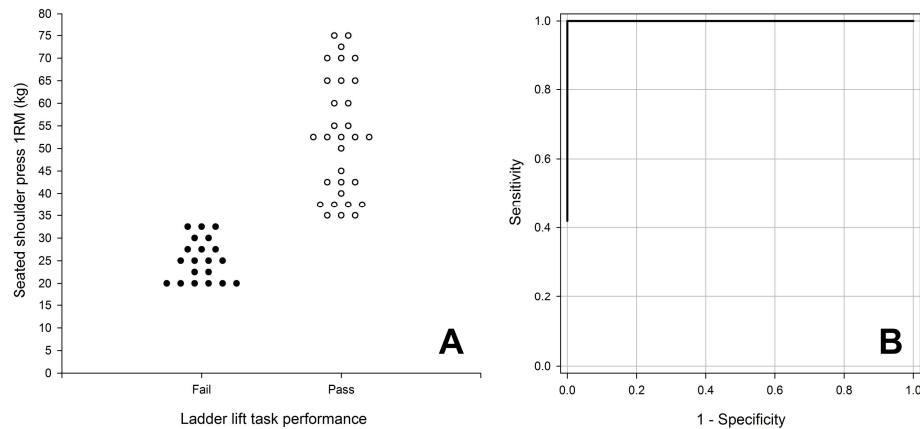


Figure 6.1. Individual performances (Pass/Fail) in the criterion ladder lift task against 1RM in the seated shoulder press test (A), and corresponding ROC curve (B) derived from these data.

On the seated shoulder press, a performance standard of 35 kg represents ideal sensitivity and specificity where both are equal (i.e. 1). At this performance level accuracy is 100%, representing a perfect predictor of criterion performance.

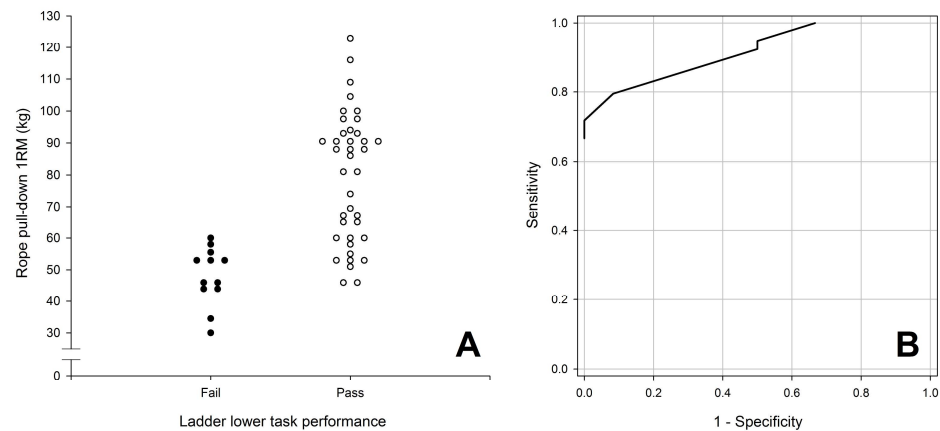


Figure 6.2. Individual performances (Pass/Fail) in the criterion ladder lower task against 1RM in seated maximal rope pull-down (A) and corresponding ROC curve (B) derived from these data.

A performance standard of 60 kg on the seated single rope pull-down test represents the closest value to ideal classification (specificity and sensitivity of 1). At this performance standard sensitivity is 0.79, specificity is 0.92 (1-specificity = 0.08), and accuracy is 82%.

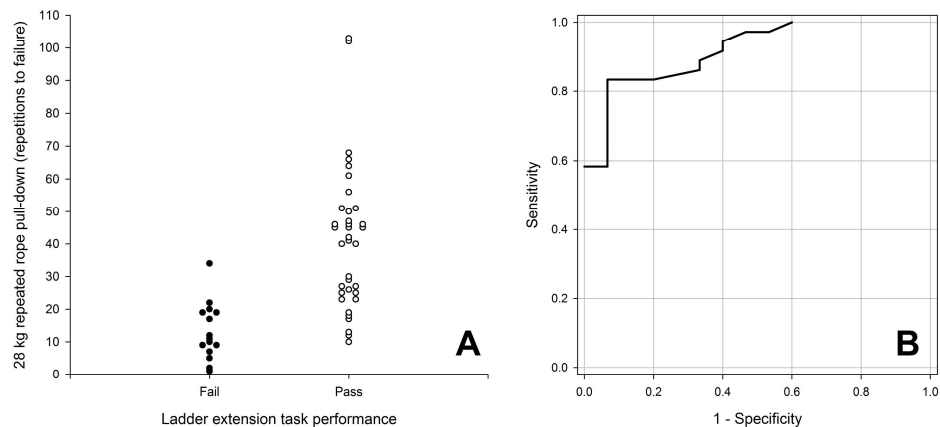


Figure 6.3. Individual performances (Pass/Fail) in criterion ladder extension task against the number of repetitions to failure during the repeated rope pull-down at 35 pulls·min⁻¹ (A) and corresponding ROC curve (B) derived from these data.

A performance standard of 23 repetitions of 28 kg on the seated repeated rope pull-down test elicited the closest value to the ideal specificity and sensitivity of 1. At this point sensitivity and specificity are 0.83 and 0.93 (1-specificity = 0.07), respectively, and accuracy is 86%.

6.4 Discussion

This study sought to assess the sensitivity and specificity of gym-based physical ability tests to predict performance in critical firefighting tasks that required the largest application of physical strength and muscular endurance. This was completed in an attempt to identify minimum muscular strength and endurance standards to ensure UK firefighters are able to perform generic tasks safely and effectively. Performance standards of 35 kg in the seated shoulder press test (surrogate for the ladder lift task), 60 kg in the seated maximal single rope pull-down test (surrogate for the ladder lower task) and 23 repetitions of 28 kg (at 35 pulls per minute) in the seated repeated rope pull-down test (surrogate for the ladder extension task) represented the optimal achievable balance of specificity and sensitivity for the respective criterion tasks. The gym-based surrogate physical ability tests and standards identified are effective at predicting the readiness of UK firefighters to perform essential occupational tasks requiring physical strength and muscular endurance.

This study applied a rigorous task analysis process which followed best practice guidelines (Tipton et al., 2012) and used highly experienced firefighters as subject matter experts to: (i) determine the critical and most arduous muscular strength and endurance tasks performed by all UK firefighters and; (ii) identify the minimum acceptable performance requirements (Stevenson et al., 2016). This ensured that the tasks identified in the research would be directly related to the critical activities of UK firefighting. Since the tasks did not require technical skill, it was possible to use civilian participants for this study. This gave a mixture of resistance trained and untrained individuals with a wide range of physical abilities. This approach likely increased the number of participants failing to complete various tasks, thus improving the predictive validity of the physical ability tests. The measures of sensitivity and specificity would have been more difficult

to determine if incumbent/trained firefighters had been recruited as participants where the vast majority of participants (if not all) could have successfully achieved all tasks.

This investigation identified that common gym-based physical ability assessments are effective at predicting performance on associated criterion tasks identified for this population, which is consistent with previous findings comparing firefighting task performance with surrogate physical ability tests (Lindberg et al., 2014; Michaelides et al., 2011; Rhea et al., 2004; Sothmann et al., 2004). However, very few of these studies identified any minimum acceptable performance standards associated with these tests. This information is a critical step for fire services when applying these surrogate tests to ensure appropriate levels of physical strength and muscular endurance for the role. The findings of this research are therefore of great benefit to fitness trainers, occupational health physicians and nurses, as well as human resource policy makers working within the UK fire and rescue services.

The only other study to identify minimum performance standards for firefighters in conjunction with muscular strength and endurance tests was conducted in a municipal fire department in the USA (Sothmann et al., 2004). A large sample of 153 serving firefighters were recruited, with 15 (10%) of those participants being female. Whilst this sample was representative of the fire department from which they were recruited, this highlights a limitation when conducting research using firefighters, since the proportion of females in the role is often relatively small. However, a noteworthy advantage of using incumbents was that the authors were able to model the impact of imposing the proposed minimum performance standards on the workforce. The authors reported that 83% of the workforce would be able to meet the minimum standards identified. Additionally, the authors indicated that the minimum cut score would identify 89% of successful performers (sensitivity) along with 72% of unsuccessful performers (specificity).

The ability to model the pass/fail rates (of any proposed standards) on the existing workforce is highly valuable to quantify possible adverse impact to specific demographic groups. However, it has to be assumed that, in terms of task performance, the sample is representative of the wider population of operational firefighters. Whilst it is well recognised that the introduction of physical employment standards may indirectly and

disproportionately affect certain demographic groups, particularly based on age or sex (Bilzon et al., 2001; Jamnik et al., 2013; Siddall et al., 2016), this information can be used to develop support mechanisms (such as physical training programmes) to minimise adverse impact to individuals and groups within an organisation. While this was not possible within the current study, it would be useful to assess the impact of implementing these standards within the UK fire and rescue services in future.

The position of each minimum performance standard in the current study was identified using ROC curves to determine the most statistically balanced combination of highest sensitivity (correctly identifying those that passed) and highest specificity (correctly identifying those that failed). This attempts to minimise the error associated with the predictive test, but typically false positives and false negatives cannot be entirely mitigated. It is possible that a standing pull-down or standing shoulder press test may have improved the likelihood of achieving higher test predictive power by closer mimicking the criterion test conditions, however this study aimed both to use standard gym based fitness equipment and to use exercises easily safeguarded by a practitioner in order to maximise the applicability of this research to fire and rescue services. In addition, while the optimum position of a standard can be determined using this analysis, there may be a requirement for a statistically determined ‘borderline’ category for tests of this nature. This would produce a lower, secondary standard and a boundary for those who may (or may not) have sufficient readiness for work. Unfortunately, there are no established guidelines for the production of these boundaries in physical employment standards research so has not been evaluated here. However, in the context of this particular study, it may be that the lowest weight increment that still maintains, for instance, 85 or 90% specificity or sensitivity could be selected as a borderline category but would depend on the specific requirements of the organisation in question, as expanded upon below.

Several limitations of this work are that without periodically implementing these tests and associated standards in the UK fire & rescue service, it is not possible to understand the true impact on the workforce or the test-retest reliability of the tests. Sensitivity and specificity are in direct opposition, where sensitivity increases, specificity decreases and vice versa. Consequently, in the likely event that a test does not achieve a

perfect predictive classification (i.e. sensitivity and specificity of 1), an organisation may choose a sub-optimal balance of these two variables. Consequently, researchers, practical end-users and/or managers would need to agree and justify the reasons for preferentially electing for higher specificity or sensitivity in a performance standard for an organisation. For example, where one might want to minimise the adverse impact on employees during a fitness test (i.e. incorrectly classifying an employee as unfit), the sensitivity of the test could be increased to reduce the possibility of this error, resulting in a lower performance standard and a higher pass rate.

However, if one felt that it was important to be extremely confident in an employee's ability to perform the task appropriately (i.e. reducing the chance of an employee incorrectly passing a fitness test), a higher specificity could be adopted, resulting in a higher performance standard and a lower pass rate. It could be that in an emergency service occupation (such as firefighting), where the impact of an employee not being able to perform the job may put lives at risk, a test that favours higher specificity may be appropriate. To the authors' knowledge, there are no globally-accepted guidelines that navigate these issues when determining physical employment standards for physically demanding or safety-critical occupations. Research focusing on repeated measures implementation of standards and tests in a workforce, and subsequent collection of impact and reliability data could help identify suitable recommendations for this, and other public safety occupations.

Cardiorespiratory performance, muscular strength and endurance are all important components of physical fitness recognised as being critical for performing firefighting duties safely and effectively. This study identified strength and muscular endurance standards on easily-replicable gym-based exercises commensurate with minimum acceptable performance requirements for essential tasks in UK firefighting. These performance standards should be applied to all UK firefighters, as part of a routine fitness assessment, to ensure that firefighters are physically able to safely carry out their work and to preserve public safety.

CHAPTER 7

General Discussion

Discussion

7.1 General Discussion

The aim of this thesis was (1) to investigate and define the minimum cardiorespiratory, strength and muscular endurance demands of undertaking UK firefighting tasks to acceptable standards and (2) to identify minimum PES and associated tests (where applicable) were to help ensure the operational effectiveness and safety of firefighting personnel working in the UK fire & rescue service. This was accomplished by building on the previous work undertaken in this field (Blacker et al., 2015; Communities and Local Government, 2009a) and by using current best practice methods for the development of PES (Tipton et al., 2012). Four novel research questions were proposed in Chapter 1, which were subsequently addressed in Chapters 3 through 6 of this work. This chapter will summarise the main findings and will discuss how the each research project has contributed to the existing knowledge in the research area.

7.1.1 What are the critical and most physically demanding generic tasks in the UK fire and rescue service?

The aim of this study was to undertake a task analysis of current UK fire and rescue activities to (a) identify the critical and most physically demanding tasks and (b) to determine the minimum acceptable performance requirements of undertaking these tasks. The main findings of this study reported that two distinct roles of UK fire and rescue personnel were classifiable in relation to the physical demands of firefighting (Stevenson et al., 2016). Those in a firefighting role (that typically undertook the most arduous roles) were performed by the ranks of Firefighter, Crew Manager and Watch Manager. Those in an incident command role (i.e. managing the fire ground) were typically performed by Station Managers and above. Eight distinct tasks (along with the minimal acceptable performance requirement) were identified as the most critical and physically demanding tasks undertaken by the generic firefighting role (Stevenson et al., 2016):

- Hose running (8 km/h)
- Casualty evacuation (6 km/h for walking and 3 km/h for rescue)
- Equipment carrying (5.5 km/h)
- Wild land fire suppression (4 km/h)
- Climbing stairs with equipment (95 steps/min)
- Lifting fire service ladders overhead (1/2 of the head of a 13.5m ladder)
- Extending fire service ladders (10.5 m ladder)
- Lowering fire service ladders (13.5 m ladder)

Two distinct tasks were identified for the incident command role (along with the corresponding minimal acceptable performance requirement):

- Walking to the scene of a wild land fire (4 km/h)
- Climbing stairs without equipment (95 steps/min)

The critical tasks performed by personnel in a firefighting role are similar to those reported by Blacker et al. (2015). Indeed, a number of the minimal acceptable performance requirements for activities involving strength and muscular endurance (lifting and extending ladders) were also similar. This suggests that many of the physical tasks undertaken by firefighters have remained relatively constant in recent years particularly in the way that firefighters use fire service ladders. This study however expanded on the previous work undertaken by Blacker et al. (2015) and has contributed to the knowledge area by establishing the minimum acceptable pace for each distinct firefighting activity. By understanding the minimum pace for each distinct firefighting task, minimum time frames for set firefighting work can be established. This may be helpful during large emergency incidents or protracted firefighting duties, where this information can be used to calculate the work capacity of firefighters, identify crewing and resource requirements to ensure that the right number of firefighters and resources are available to complete the task and to protect firefighters from over-exertion. For the purposes of this work, this information was sought to investigate the physical demand of performing these separate tasks in order to establish minimum cardiorespiratory standards.

7.1.2 What are the role-related minimum cardiorespiratory fitness standards for firefighters and commanders?

In order to establish the minimum cardiorespiratory fitness standards for firefighters and incident commanders, it was first necessary to establish the physical demands of performing these critical tasks to the minimum acceptable performance requirement. From this, the physiological data was used to derive a minimum cardiorespiratory fitness standard for personnel in firefighting and incident command roles. The main findings identified that the peak metabolic demands for firefighting roles ranged from 29 ml·kg·min⁻¹ to 47 ml·kg·min⁻¹ and from 23 ml·kg·min⁻¹ to 35 ml·kg·min⁻¹ for incident command roles. From this data minimum cardiorespiratory fitness standards of 42.3 ml·kg·min⁻¹ and 36.8 ml·kg·min⁻¹ were derived for those undertaking firefighting and incident command roles respectively.

The recommended cardiorespiratory fitness standard of 42.3 ml·kg·min⁻¹ in this study is similar to both unpublished (Stevenson et al., 2009) and published (Bilzon et al., 2001; Gledhill & Jamnik, 1992a; Jamnik et al., 2013; Lemon & Hermiston, 1977; National Fire Protection Association, 2007) recommendations in other firefighting populations. One of the strengths of this study in establishing the firefighter fitness standard was that it compared the two most common best-practice methods used to derive a fitness standard from physical demands data. Whilst the method used in this current study was adopted from the work of Bilzon et al. (2001) eliciting a fitness standard of 42.3 ml·kg·min⁻¹, a very similar fitness standard of 41.9 ml·kg·min⁻¹ was calculated using the work of Jamnik et al. (2013). Since the derived standard from this study is so similar to both the Jamnik et al. (2013) calculation as well as other recommendations for cardiorespiratory fitness standards, this strengthened the reasoning for the minimum cardiorespiratory fitness standard described for UK firefighters in this work.

For the first time in the UK this study also objectively quantified the minimum cardiorespiratory fitness standards for incident commanders. Prior to this work, little evidence was available to recommend a minimum standard for those in an incident management role, which may have led services and individuals to play down the need for incident commanders to remain physically fit. Indeed a Government review into the

effects of aging on physical fitness in 2012 indicated that the minimum cardiorespiratory fitness requirement to perform this type of role would be approximately $25 \text{ ml}\cdot\text{kg}\cdot\text{min}^{-1}$ (Williams, Wilkinson, Richmond, & Rayson, 2012). The evidence from this current study has however indicated that climbing stairs in fire kit whilst wearing breathing apparatus elicits a greater physical strain on incident commanders than previously thought, with individuals requiring a minimum cardiorespiratory fitness standard of $36.8 \text{ ml}\cdot\text{kg}\cdot\text{min}^{-1}$ (Siddall et al., 2016). This evidence delivers clear guidance to fire and rescue services of the importance of maintaining fitness levels for managers even if they are not directly involved in the most demanding of firefighting tasks (Siddall et al., 2016; Stevenson et al., 2016). Considering managers in the fire and rescue service may work into their sixties, maintaining levels of physical fitness will help them sustain their health for longer, minimise the risk of injury and sickness. There is some evidence that remaining fit may even help with decision-making and the mental demands of managing an emergency incident (Suominen-Troyer, Davis, Ismail, & Salvendy, 1986; Throne, Bartholomew, Craig, & Farrar, 2000) thus improving their own safety and the safety of others who they are instructing.

7.1.3 What is the validity and reliability of a firefighting simulation test (FFST)?

As well as recommending a number of gym-based physical fitness tests, a FFST was designed and investigated to see if it could be appropriately used as a valid and reliable criterion test of operational fitness. Therefore the aim of this study was to assess the validity of the FFST to estimate maximal cardiorespiratory fitness (VO_2max) and establish the test-retest reliability. The main findings from this study indicated that the time taken to complete the FFST demonstrated a strong inverse correlation with VO_2max demonstrating that performance on the FFST is strongly determined by cardiorespiratory fitness. Furthermore, the repeatability of the simulation was highly reliable ($r = 0.842$). However, the error associated with the simulation (SEE of $6.13 \text{ ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$), and the test-retest coefficient of variation of 4.5% suggests it may not be suitably accurate enough to predict VO_2max and therefore be used as a stand-alone fitness test.

Whilst research investigating the association between maximal firefighting performance and maximum cardiorespiratory fitness is not new (Boyd et al., 2015; Rhea et al., 2004; Von Heimburg et al., 2013; Williams-Bell et al., 2009; Williford et al., 1999), a number of methodological differences in the previous work have made it difficult to directly compare studies and the association between firefighting performance and cardiorespiratory fitness. Some studies used set periods of walking between firefighting tasks that may not be representative of a real fire scenario and allowed for small amounts of rest to be achieved during the firefighting task (Williams-Bell et al., 2009). Other studies may not have completed the tasks as fast as possible being instructed to complete the tasks “with no unnecessary waste of time” thus potentially affecting the relationship between task performance and the fitness parameter. In order to accurately assess the relationship between cardiorespiratory fitness and firefighting performance, this study established a simulation using critical and physically demanding UK firefighting tasks (Stevenson et al., 2016) in a realistic fire simulation, with participants completing the task as fast as possible with those reporting an RPE of less than 17 (i.e. not maximal performance) being removed from the analysis. Thus this study used a rigorous methodology to ensure that both the firefighting task and the cardiorespiratory fitness test were completed maximally thus contributing to the knowledge area of this work.

7.14 What are the minimum muscular strength and endurance requirements for firefighters?

In order to recommend strength and muscular endurance standards for firefighting roles, it was necessary to assess the sensitivity and specificity of common and replicable gym-based physical ability tests to predict performance on the critical and most physically demanding firefighting tasks that require the largest application of physical strength and muscular endurance. The main findings from the study indicated that all of the gym-based physical ability tests based described were effective at predicting effective firefighting performance and that performance standards of 35 kg in the seated shoulder press test (surrogate for the ladder lift task), 60 kg in the seated maximal single rope pull-down test (surrogate for the ladder lower task), and 23 repetitions of 28 kg (at 35 pulls per minute) in the seated repeated rope pull-down test (surrogate for the ladder extension task)

represented the optimal achievable balance of specificity and sensitivity for the respective criterion tasks.

The association of both strength and muscular endurance performance with successful firefighting ability is well reported in the research (Henderson et al., 2007; Lindberg et al., 2014; Michaelides et al., 2011; Rhea et al., 2004; Von Heimburg et al., 2006). Indeed much of this research identifies that surrogate tests can be used to predict firefighting performance. However, only one study has recommended minimum strength and muscular endurance standards for firefighters (Sothmann et al., 2004), which seems surprising considering the importance of these physical attributes. A possible explanation for the lack of minimum standards could be that the studies mentioned here did not have clearly identified minimum performance requirements for firefighting activities investigated, which would subsequently make it difficult to identify minimum PES. The findings from this study therefore contribute greatly to the knowledge area and may be applicable to other fire services that use fire service ladders of similar weights or those wishing to develop their own.

This study also raised an important issue regarding balancing test sensitivity and specificity to identify minimum performance standards. In this study an optimal balance of sensitivity and specificity was chosen as there were to the author's knowledge no accepted guidelines for interpreting levels of sensitivity and specificity to identify a PES. Whilst a number of factors may affect decisions taken on the balance of sensitivity and specificity chosen, one could make a strong case for higher levels of specificity in emergency service or public safety roles. In this example, one might wish to see a test sensitivity of at least 0.9 to ensure an incumbent has the necessary physical ability to perform the required tasks / work. As international consensus has been recently been reached in the steps needed to develop a robust PES, discussion and agreement on the minimum requirements for test sensitivity and / or specificity for different types of PES may allow other researchers to develop physical standards in demanding occupations thus improving the safety of employees and the public.

7.15 Practical implications

This project was initiated to answer a number of specific research questions raised by the Chief Fire Officers Association (CFOA) on behalf of the UK fire and rescue service. The minimum acceptable performance requirements (Chapter 3), cardiorespiratory fitness standards (Chapter 4), FFST (Chapter 5) and strength and muscular endurance standards (Chapter 6) have all accepted by CFOA and published as national guidance to all UK fire and rescue services making this project highly practical to industry. This guidance was published within a fitness management process to ensure the efficient but safe management of fire and rescue service personnel (Figure 7.1). Finally, the task analysis methodology (Chapter 3) was also developed and published in a way that it could be applied in other physically demanding industries wishing to undertake a task analysis thus also making it highly practical and adaptable.

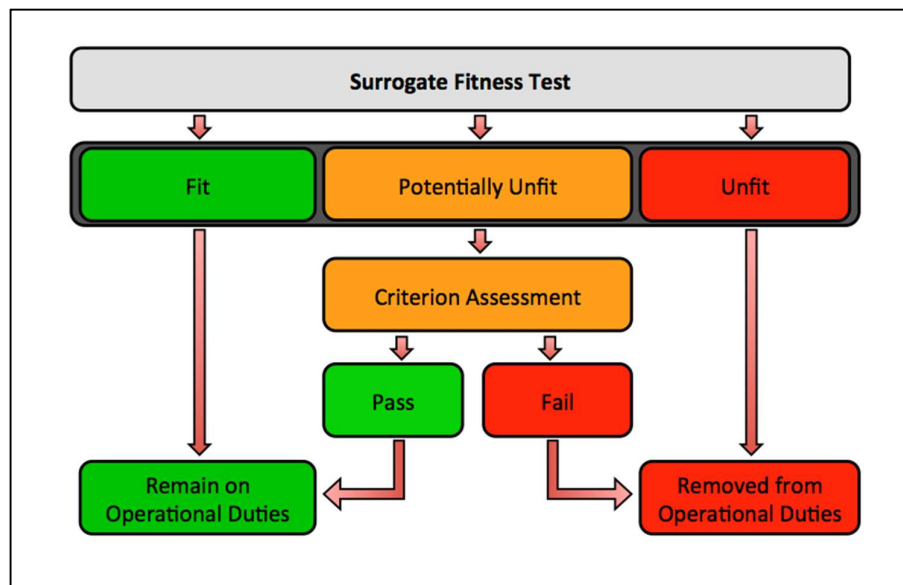


Figure 7.1 Fitness management process for the UK fire and rescue services

7.16 Supporting research

During the writing of this thesis, a number of other research projects have contributed to this area of study, which compliments the work presented here. In 2015 Blacklock et al. developed a systematic task analysis process using subject matter experts using a common

and essential task in the Canadian Armed Forces as a practical example (Blacklock, Reilly, Spivock, Newton, & Olinek, 2015). The study identified a way in which a realistic simulation and minimum performance standard could be identified. This study would have been helpful in the writing of Chapter 3 where the issue of determining the minimum acceptable pace was identified as being an important step in the research but lacked practical steps to achieve this.

In 2016, 2 studies provided further insights into the effect of human load carriage. Peoples et al. investigated the effect of load carriage on work duration during progressive treadmill walking (Peoples, Lee, Notley, & Taylor, 2016). This work identified that thoracic loading significantly reduced exercise tolerance and acceptable work durations were less than derived from previously published methods. A review article also suggested that existing unloaded workload guidelines are inappropriate for tasks involving load carrying (Drain, Billing, Neesham-Smith, & Aisbett, 2016). These articles would be of importance particularly in Chapter 4 where maximal permissible work rates were calculated from previous scientific evidence.

7.2 Future Directions

7.2.1 Specialist roles

Understanding the cardiorespiratory, strength and muscular demands of generic UK firefighting tasks and establishing role related PES is critical to help ensure the health, fitness and operational performance of all UK firefighting personnel. However, firefighters also undertake a number of specialist roles depending on the geography of their work location (fire station). These include activities such as working road traffic collisions, water or mud rescue and Urban Search and Rescue (USAR) where UK firefighters may be required to attend both domestic and international large-scale natural and made disasters such as earthquakes, flooding, chemical spills and terrorist attacks. These activities were identified in the first study (Chapter 3) as physically demanding but not applicable to all UK firefighters and as such were removed from the task analysis and not investigated in this work. However, investigating these specialist roles and establishing whether their physical demands are different to the generic firefighting role

is important to ensure that fitness levels are both role related and task specific thus ensuring the operational effectiveness and safety of firefighters.

7.2.2 Trainability of UK Firefighters

The majority of research into the physicality of firefighting since the turn of the century has focused on PES for firefighter applicants (Blacker et al., 2015) to ensure legally defensible employment test and standards. More recently, the work identified in this project has focused on PES to ensure the safety and operational effectiveness of incumbents (Siddall et al., 2016; Stevenson et al., 2016, 2017). Whilst it is important to ensure that both applicants and incumbents can demonstrate the necessary levels of fitness for safe and effective firefighting, it is well known that levels of physical fitness can fluctuate in a short space of time (American College of Sports Medicine, 2010). Indeed, a number of factors can greatly affect levels of physical fitness including age (Haisman, 1996; Kenny, Groeller, McGinn, & Flouris, 2016; Shephard, 1999; Williams et al., 2012) body composition and weight (Baur, Christophi, & Kales, 2012; Baur, Christophi, Tsismenakis, et al., 2012) and gender (Roberts, Gebhardt, Gaskill, Roy, & Sharp, 2016). Considering the transient nature of physical fitness and the impact this can have on firefighter safety, employee wellbeing and career longevity, very little research is directed at strategies to help employees maintain levels of physical fitness throughout a firefighter's career. This is despite a government report indicating that 'physical fitness offers a cost-effective method of enhancing [firefighting] performance and improving employee health' (Office of the Deputy Prime Minister, 2004b). Perhaps one reason why this has not been undertaken thus far is because until now the physical demands of UK firefighting has remained unknown. Since this work is now complete, it would seem prudent that further research be undertaken to develop appropriate physical training regimes that help firefighters maintain appropriate levels of fitness for their role(s) in order to minimise the detrimental effects of declines in physical fitness such as increased injury risk, periods of non-operational duties or early retirement due to ill health.

7.2.3 The effects of physical fitness on decision making and mental ability in firefighters

The findings from the body of work presented here add to the recognition of the importance of physical fitness for successful firefighting performance in the UK fire and rescue service. However, whilst the importance of physical fitness on physical performance is well recognised, the importance of physical conditioning on mental ability is less well understood. However, it is recognised that both firefighters and managers are required to make important split second decisions either during emergency response driving or at the scene of an emergency incident, which can greatly affect the safety of workers and the public. Whilst there is some evidence to indicate that improvements in physical fitness can improve decisions making (Suominen-Troyer et al., 1986) and reduce the psychological stress during strategy and tactical simulations in firefighters (Throne et al., 2000), the evidence base is relatively weak. Considering the importance of making correct decisions, it would be of interest to expand on the current research available and investigate whether improving physical fitness affects decision making in UK firefighters and if so whether altering factors such as the intensity, duration or type of physical activity has an effect on mental performance in firefighters and managers. This information coupled with the detailed training plan to cope with the physical demands of the work would ensure a comprehensive, evidence based physical fitness programme to help maintain both the physical and mental fitness of firefighters and managers.

References

- Adams, E. M. (2016). Human rights at work: Physical standards for employment and human rights law. *Appl Physiol Nutr Metab*, 41(6 Suppl 2), S63-73.
- American College of Sports Medicine. (2010). *ACSM's Resource Manual for Guidelines for Exercise Testing and Prescription*. 6th ed.
- Arngrimsson, S. A., Stewart, D. J., Borrani, F., Skinner, K. A., & Cureton, K. J. (2003). Relation of heart rate to percent VO₂ peak during submaximal exercise in the heat. *J Appl Physiol* (1985), 94(3), 1162-1168.
- Astrand, P.O., & Rodahl, K. (1970). *The text book of work physiology* (Vol. 1st edn.).
- Barnekow-Bergkvist, M., Aasa, U., Angquist, K. A., & Johansson, H. (2004). Prediction of development of fatigue during a simulated ambulance work task from physical performance tests. *Ergonomics*, 47(11), 1238-1250.
- Baur, D. M., Christophi, C. A., & Kales, S. N. (2012). Metabolic syndrome is inversely related to cardiorespiratory fitness in male career firefighters. *J Strength Cond Res*, 26(9), 2331-2337.
- Baur, D. M., Christophi, C. A., Tsismenakis, A. J., Jahnke, S. A., & Kales, S. N. (2012). Weight- perception in male career firefighters and its association with cardiovascular risk factors. *BMC Public Health*, 12, 480.
- Billat, L. V., & Koralsztein, J. P. (1996). Significance of the velocity at VO₂max and time to exhaustion at this velocity. *Sports Med*, 22(2), 90-108.
- Bilzon, J. L., Scarpello, E. G., Bilzon, E., & Allsopp, A. J. (2002). Generic task-related occupational requirements for Royal Naval personnel. *Occup Med (Lond)*, 52(8), 503-510.

- Bilzon, J. L., Scarpello, E. G., Smith, C. V., Ravenhill, N. A., & Rayson, M. P. (2001). Characterization of the metabolic demands of simulated shipboard Royal Navy fire-fighting tasks. *Ergonomics*, 44(8), 766-780.
- Bink, B. (1962). The physical working capacity in relation to working time and age. *Ergonomics*, 5(1), 25-28.
- Blacker, S. D., Rayson, M. P., Wilkinson, D. M., Carter, J. M., Nevill, A. M., & Richmond, V. L. (2015). Physical employment standards for UK fire and rescue service personnel. *Occup Med (Lond)*.
- Blacker, S. D., Wilkinson, D. M., Bilzon, J. L., & Rayson, M. P. (2008). Risk factors for training injuries among British Army recruits. *Mil Med*, 173(3), 278-286.
- Blacklock, R. E., Reilly, T. J., Spivock, M., Newton, P. S., & Olinek, S. M. (2015). Standard Establishment Through Scenarios (SETS): A new technique for occupational fitness standards. *Work*, 52(2), 375-383.
- Blondel, N., Berthoin, S., Billat, V., & Lensel, G. (2001). Relationship between run times to exhaustion at 90, 100, 120, and 140% of $\dot{V}O_{2\max}$ and velocity expressed relatively to critical velocity and maximal velocity. *Int J Sports Med*, 22(1), 27-33.
- Borg, G. A. (1982). Psychophysical bases of perceived exertion. *Med Sci Sports Exerc*, 14(5), 377-381.
- Bos, J., Mol, E., Visser, B., & Frings-Dresen, M. (2004). The physical demands upon (Dutch) fire-fighters in relation to the maximum acceptable energetic workload. *Ergonomics*, 47(4), 446-460.
- Boyd, L., Rogers, T., Docherty, D., & Petersen, S. (2015). Variability in performance on a work simulation test of physical fitness for firefighters. *Appl Physiol Nutr Metab*, 40(4), 364-370.

- Buckley, J. P., Sim, J., Eston, R. G., Hession, R., & Fox, R. (2004). Reliability and validity of measures taken during the Chester step test to predict aerobic power and to prescribe aerobic exercise. *Br J Sports Med*, 38(2), 197-205.
- Cartledge, P. (2004). *The Spartans: The World of the Warrior-Heroes of Ancient Greece. Reprint edition (August 10, 2004)*: Vintage Books, New York.
- Communities and Local Government. (2009a). National firefighter selection process. Development and validation of national firefighter selection tests: physical tests. *Communities and Local Government Publications, UK Government*.
- Constable, S., & Palmer, B. (2000). *The Process of Physical Fitness Standards Development*.
- Davis, P. O., Dotson, C. O., & Maria, D. L. S. (1982). Relationship between simulated fire fighting tasks and physical performance-measures. *Med Sci Sports Exerc*, 14(1), 65-71.
- Drain, J., Billing, D., Neesham-Smith, D., & Aisbett, B. (2016). Predicting physiological capacity of human load carriage - a review. *Appl Ergon*, 52, 85-94.
- Dreger, R. W., & Petersen, S. R. (2007). Oxygen cost of the CF-DND fire fit test in males and females. *Appl Physiol Nutr Metab*, 32(3), 454-462.
- Duffield, R., Dawson, B., Pinnington, H. C., & Wong, P. (2004). Accuracy and reliability of a Cosmed K4b2 portable gas analysis system. *J Sci Med Sport*, 7(1), 11-22.
- East, W.B. (2013). *A Historical Review and Analysis of Army Physical Readiness Training and Assessment. Combat Studies Institute Press - US Army Combined Arms Center. Fort Leavenworth, Kansas*.
- Eglin, C. M., Coles, S., & Tipton, M. J. (2004). Physiological responses of fire-fighter instructors during training exercises. *Ergonomics*, 47(5), 483-494.

- Elsner, K. L., Kolkhorst, F. W. (2008). Metabolic demands of simulated firefighting tasks. *Ergonomics*, 51(9), 1418-1425.
- Epstein, Y., Yanovich, R., Moran, D. S., & Heled, Y. (2013). Physiological employment standards IV: integration of women in combat units physiological and medical considerations. *Eur J Appl Physiol*, 113(11), 2673-2690.
- FireFit Steering Group. (2009). *Survey of current fitness practices*.
- Gamble, R. P., Stevens, A. B., McBrien, H., Black, A., Cran, G. W., & Boreham, C. A. (1991). Physical fitness and occupational demands of the Belfast ambulance service. *Br J Ind Med*, 48(9), 592-596.
- Gledhill, N., & Jamnik, V. K. (1992a). Characterization of the physical demands of firefighting. *Canadian Journal of Sport Sciences-Revue Canadienne Des Sciences Du Sport*, 17(3), 207-213.
- Gledhill, N., & Jamnik, V. K. (1992b). Development and validation of a fitness screening protocol for firefighter applicants. *Canadian Journal of Sport Sciences-Revue Canadienne Des Sciences Du Sport*, 17(3), 199-206.
- Gleser, M. A., & Vogel, J. A. (1973). Endurance capacity for prolonged exercise on the bicycle ergometer. *J Appl Physiol*, 34(4), 438-442.
- Goheer, A., Bailey, M., Gittelsohn, J., & Pollack, K. M. (2013). Fighting fires and fat: An intervention to address obesity in the fire service. *J Nutr Educ Behav*.
- Great Britain Parliament. (1974). Health & Safety at Work Act. *London: HSMO*.
- Great Britain Parliament. (2010). Equality Act. *London: HSMO*.
- Gumieniak, Robert J., Jamnik, Veronica K., & Gledhill, Norman. (2013). Catalog of canadian fitness screening protocols for public safety occupations that qualify as

a bona fide occupational requirement. *Journal of Strength and Conditioning Research*, 27(4), 1168-1173.

Haisman, M. (1996). *Age Limit of Serving Firefighters*.

Health & Safety Executive. (2004). *Manual Handling Operations Regulations 1992 (as amended) Guidance on Regulations*. (Third Edition (2004)).

Health & Safety Executive. (2010). *Striking the balance between operational and health and safety duties in the Fire and Rescue Service*
<http://www.hse.gov.uk/services/fire/duties.pdf>.

Henderson, Norman D., Berry, Michael W., & Matic, Tomislov. (2007). Field measures of strength and fitness predict firefighter performance on physically demanding tasks. *Personnel Psychology*, 60(2), 431-473.

Her Majesty's Fire Service Inspectorate. (2000). *Fit For Duty? Seeking a Healthier Fire Service*.

Her Majesty's Fire Service Inspectorate. (2004). *Fire Service Manual - Volume 4: Fire Service Training* TSO (The Stationery Office).

Holmer, I., & Gavhed, D. (2007). Classification of metabolic and respiratory demands in fire fighting activity with extreme workloads. *Appl Ergon*, 38(1), 45-52.

Home Office. (2016). *Fire Statistics Monitor: April 2015 to March 2016. Statistical Bulletin 09/16*.

Howe, C. C., Matzko, R. O., Piasek, F., Pitsiladis, Y. P., & Easton, C. (2014). Stability of the K4b(2) portable metabolic analyser during rest, walking and running. *J Sports Sci*, 32(2), 157-163.

- Industrial Organisational Solutions. (2015). Job Analysis Study Technical Report For the Position of Career and Volunteer Firefighter in the State of Florida: Division of State Fire Marshall, Bureau of Fire Standards and Training.
- Jamnik, V.K., Gumienak, R., & Gledhill, N. (2010). Applying the Meiorin Decision requirements to the fitness test for correctional officer applicants; examining adverse impact and accommodation. *Applied Physiology Nutrition and Metabolism*, 35(1), 71-81.
- Jamnik, V.K., Gumienak, R., & Gledhill, N. (2013). Developing legally defensible physiological employment standards for prominent physically demanding public safety occupations: a Canadian perspective. *Eur J Appl Physiol*.
- Jamnik, V.K., Thomas, Scott G., Burr, Jamie F., & Gledhill, Norman. (2010). Construction, validation, and derivation of performance standards for a fitness test for correctional officer applicants. *Applied Physiology Nutrition and Metabolism-Physiologie Appliquee Nutrition Et Metabolisme*, 35(1), 59-70.
- Jamnik, V.K., Thomas, Scott G., Shaw, Jim A., & Gledhill, Norman. (2010). Identification and characterization of the critical physically demanding tasks encountered by correctional officers. *Applied Physiology Nutrition and Metabolism*, 35(1), 45-58.
- Kales, Stefanos N., Soteriades, Elpidoforos S., Christophi, Costas A., & Christiani, David C. (2007). Emergency duties and deaths from heart disease among firefighters in the United States. *New England Journal of Medicine*, 356(12), 1207-1215.
- Kane, M. (1994). Validating the performance standards associated with passing scores. *Review of Educational Research*, 64, 425-461.
- Kenny, G. P., Groeller, H., McGinn, R., & Flouris, A. D. (2016). Age, human performance, and physical employment standards. *Appl Physiol Nutr Metab*, 41(6 Suppl 2), S92-S107.

- Kilbom, A. (1980). Physical work capacity of firemen. With special reference to demands during fire fighting. *Scand J Work Environ Health*, 6(1), 48-57.
- Knapik, J. (1989). The Army Physical Fitness Test (APFT): a review of the literature. *Mil Med*, 154(6), 326-329.
- Knapik, J. J., & East, W. B. (2014). History of United States Army physical fitness and physical readiness training. *US Army Med Dep J*, 5-19.
- Labour Resarch Department. (2008). The fire brigades union. In the line of duty. Firefighter deaths in the uk since 1978. *Labour Resarch Department*.
- Lemon, P. W. R., & Hermiston, R. T. (1977). Human energy-cost of fire fighting. *Journal of Occupational and Environmental Medicine*, 19(8), 558-562.
- Lewis, D. M., Mitzel, H. C., Green, D. R., & Patz, R. J. (1999). The Bookmark standard setting procedure. *Monterey, CA: McGraw-Hill*.
- Lindberg, A. S., Oksa, J., & Malm, C. (2014). Laboratory or field tests for evaluating firefighters' work capacity? *PLoS One*, 9(3), e91215.
- Louhevaara, V., Smolander, J., Korhonen, O., & Tuomi, T. (1986). Maximal working times with a self-contained breathing apparatus. *Ergonomics*, 29(1), 77-85.
- McClave, S. A., & Snider, H. L. (1992). Use of indirect calorimetry in clinical nutrition. *Nutr Clin Pract*, 7(5), 207-221.
- Michaelides, M. A., Parpa, K. M., Henry, L. J., Thompson, G. B., & Brown, B. S. (2011). Assessment of physical fitness aspects and their relationship to firefighters' job abilities. *J Strength Cond Res*, 25(4), 956-965.

- Milligan, G. S., Reilly, T. J., Zumbo, B. D., & Tipton, M. J. (2016). Validity and reliability of physical employment standards. *Appl Physiol Nutr Metab*, 41(6 Suppl 2), S83-91.
- Milligan, Gemma S., & Tipton, Michael J. (2013). A report on the first conference on physiological and physical employment standards, Canberra, Australia November 27 to 28, 2012. *Extreme physiology & medicine*, 2(1), 14-14.
- Munir, F., Clemes, S., Houdmont, J., & Randall, R. (2012). Overweight and obesity in UK firefighters. *Occup Med (Lond)*, 62(5), 362-365.
- National Fire Protection Association. (2007). NFPA Association 1582: Standard on Comprehensive Occupational Medical Program for Fire Departments. Quincy, MA: National Fire Protection Association, 2007.
- Nielsen, B., & Nybo, L. (2003). Cerebral changes during exercise in the heat. *Sports Med*, 33(1), 1-11.
- Nybo, L. (2007). Exercise and heat stress: cerebral challenges and consequences. *Prog Brain Res*, 162, 29-43.
- Office of the Deputy Prime Minister. (2000). *Fit for duty ? Seeking a healthier fire service. A thematic review of sickness absence and ill-health retirements in the fire service in england, wales and northern ireland.*
- Office of the Deputy Prime Minister. (2003). *Our Fire & Rescue Service.*
- Office of the Deputy Prime Minister. (2004a). *Medical and occupational evidence for recruitment and retention in the Fire and Rescue Service.*
- Office of the Deputy Prime Minister. (2004b). *Operational physiological capabilities of firefighters: Literature review and research recommendations.*

- Payne, W., & Harvey, J. (2010). A framework for the design and development of physical employment tests and standards. *Ergonomics*, 53(7), 858-871.
- Peoples, G. E., Lee, D. S., Notley, S. R., & Taylor, N. A. (2016). The effects of thoracic load carriage on maximal ambulatory work tolerance and acceptable work durations. *Eur J Appl Physiol*, 116(3), 635-646.
- Perroni, Fabrizio, Tessitore, Antonio, Cortis, Cristina, Lupo, Corrado, D'Artibale, Emanuele, Cignitti, Lamberto, & Capranica, Laura. (2010). Energy cost and energy sources during a simulated firefighting activity. *Journal of Strength and Conditioning Research*, 24(12), 3457-3463.
- Petersen, S. R., Anderson, G. S., Tipton, M. J., Docherty, D., Graham, T. E., Sharkey, B. J., & Taylor, N. A. (2016). Towards best practice in physical and physiological employment standards. *Appl Physiol Nutr Metab*, 41(6 Suppl 2), S47-62.
- Pinnington, H. C., Wong, P., Tay, J., Green, D., & Dawson, B. (2001). The level of accuracy and agreement in measures of FEO₂, FECO₂ and VE between the Cosmed K4b2 portable, respiratory gas analysis system and a metabolic cart. *J Sci Med Sport*, 4(3), 324-335.
- Plat, M. J., Frings-Dresen, M. H., & Sluiter, J. K. (2011). A systematic review of job-specific workers' health surveillance activities for fire-fighting, ambulance, police and military personnel. *Int Arch Occup Environ Health*, 84(8), 839-857.
- Rayson, M. P. (2000). Fitness for work: the need for conducting a job analysis. *Occupational Medicine-Oxford*, 50(6), 434-436.
- Rayson, M. P. (2000a). *Job Analysis*. In: Constable, S. & Palmer, B. ed. *The Process of Physical Fitness Standards Development*. pp67-98.

- Rayson, M. P., Holliman, D., & Belyavin, A. (2000). Development of physical selection procedures for the British Army. Phase 2: relationship between physical performance tests and criterion tasks. *Ergonomics*, 43(1), 73-105.
- Rayson, M.P, Wilkinson, D.M, Carter, J.M., & Nevill, A.M. (2009a). *National Firefighter Selection Process Development and Validation of National Firefighter Selection Tests: Physical Tests*. London, UK: Department for Communities and Local Government.
- Reilly, T., Iggleden, C., Gennser, M., & Tipton, M. (2006). Occupational fitness standards for beach lifeguards. Phase 2: the development of an easily administered fitness test. *Occup Med (Lond)*, 56(1), 12-17.
- Reilly, T., Wooller, A., & Tipton, M. (2006). Occupational fitness standards for beach lifeguards. Phase 1: the physiological demands of beach lifeguarding. *Occup Med (Lond)*, 56(1), 6-11.
- Rhea, M. R., Alvar, B. A., & Gray, R. (2004). Physical fitness and job performance of firefighters. *J Strength Cond Res*, 18(2), 348-352.
- Richmond, V. L., Rayson, M. P., Wilkinson, D. M., Carter, J. M., & Blacker, S. D. (2008). Physical demands of firefighter search and rescue in ambient environmental conditions. *Ergonomics*, 51(7), 1023-1031.
- Roberts, D., Gebhardt, D. L., Gaskill, S. E., Roy, T. C., & Sharp, M. A. (2016). Current considerations related to physiological differences between the sexes and physical employment standards. *Appl Physiol Nutr Metab*, 41(6 Suppl 2), S108-120. doi: 10.1139/apnm-2015-0540
- Roberts, M. A., O'Dea, J., Boyce, A., & Mannix, E. T. (2002). Fitness levels of firefighter recruits before and after a supervised exercise training program. *Journal of Strength and Conditioning Research*, 16(2), 271-277.
- Roberts, S. (2002). Hazardous occupations in Great Britain. *Lancet*, 360(9332), 543-544.

- Rogers, T., Docherty, D., & Petersen, S. (2014). Establishment of performance standards and a cut-score for the Canadian Forces firefighter physical fitness maintenance evaluation (FF PFME). *Ergonomics*, 57(11), 1750-1759.
- Schrack, J. A., Simonsick, E. M., & Ferrucci, L. (2010). Comparison of the Cosmed K4b(2) portable metabolic system in measuring steady-state walking energy expenditure. *PLoS One*, 5(2), e9292.
- Scott, G.E. (1988). *The Physical Fitness of Firemen*. Home Office Scientific Research and Development Branch.
- Sheaff, Andrew K., Bennett, Angela, Hanson, Erik D., Kim, You-Sin, Hsu, Jeffrey, Shim, Jae K., . . . Hurley, Ben F. (2010). Physiological determinants of the candidate physical ability test in firefighters. *J Strength Cond Res*, 24(11), 3112-3122.
- Shephard, R. J. (1955). A critical examination of the Douglas bag technique. *J Physiol*, 127(3), 515-524.
- Shephard, R. J. (1999). Age and physical work capacity. *Exp Aging Res*, 25(4), 331-343.
- Siddall, A. G., Stevenson, R. D., Turner, P. F., Stokes, K. A., & Bilzon, J. L. (2016). Development of role-related minimum cardiorespiratory fitness standards for firefighters and commanders. *Ergonomics*, 59(10), 1335-1343.
- Smith, D. L., Barr, D. A., & Kales, S. N. (2013). Extreme sacrifice: sudden cardiac death in the US Fire Service. *Extrem Physiol Med*, 2(1), 6.
- Soteriades, Elpidoforos S., Smith, Denise L., Tsismenakis, Antonios J., Baur, Dorothee M., & Kales, Stefanos N. (2011). Cardiovascular disease in us firefighters. A systematic review. *Cardiology in Review*, 19(4), 202-215.

- Sothmann, M. S., Landy, F., & Saupe, K. (1992). Age as a bona-fide occupational qualification for firefighting - a review on the importance of measuring aerobic power. *Journal of Occupational and Environmental Medicine*, 34(1), 26-33.
- Sothmann, M.S., Gebhardt, D. L., Baker, T. A., Castello, G. M., & Sheppard, V. A. (2004). Performance requirements of physically strenuous occupations: validating minimum standards for muscular strength and endurance. *Ergonomics*, 47(8), 864-875.
- Sothmann, M.S., Saupe, K.W., Jasenof, D., Blaney, J., Donaghue Fuhrman, S., & Woulfe, T. (1990). Advancing age and the cardiorespiratory stress of fire suppression: Determining a minimum standard aerobic fitness. *Human Performance*, 3(4), 217-236.
- Stevenson, R. D., Siddall, A. G., Turner, P. F., & Bilzon, J. L. (2016). A Task Analysis Methodology for the Development of Minimum Physical Employment Standards. *J Occup Environ Med*, 58(8), 846-851.
- Stevenson, R. D., Siddall, A. G., Turner, P. F., & Bilzon, J. L. (2017). Physical Employment Standards for UK Firefighters: Minimum Muscular Strength and Endurance Requirements. *J Occup Environ Med*, 59(1), 74-79.
- Stevenson, R.D.M., Wilsher, P. & , & Sykes, K. (2009). Fitness for fire & rescue. Standards, protocol and policy.
- Stout, S.E. (1921). Training Soldiers for the Roman Legion *The Classical Journal*, 16 (7), 423-431.
- Suominen-Troyer, S., Davis, K. J., Ismail, A. H., & Salvendy, G. (1986). Impact of physical fitness on strategy development in decision-making tasks. *Percept Mot Skills*, 62(1), 71-77.

- Taylor, N. A., Fullagar, H. H., Mott, B. J., Sampson, J. A., & Groeller, H. (2015). Employment Standards for Australian Urban Firefighters: Part 1: The Essential, Physically Demanding Tasks. *J Occup Environ Med*, 57(10), 1063-1071.
- Taylor, N. A., & Groeller, H. (2003). Work-based physiological assessment of physically-demanding trades: a methodological overview. *J Physiol Anthropol Appl Human Sci*, 22(2), 73-81.
- The Chief Fire Officers Association. (2012). Responsibilities for Fitness Testing - Circular 2012-012.
- The Firefighter Foundation. (2017). Retrieved 15/02/2017, from <http://www.firefighterfoundation.org.uk>
- Throne, L.C, Bartholomew, J.B. , Craig, J, & Farrar, R.P. (2000). Stress Reactivity in Fire Fighters: An Exercise Intervention. *International Journal of Stress Management*, 7(4), 235-246.
- Tipton, M. J., Milligan, G. S., & Reilly, T. J. (2012). Physiological employment standards I. Occupational fitness standards: objectively subjective? *Eur J Appl Physiol*.
- Truxillo, D.M., Paronto, M.E., Collins, M. & Sulzer, J.L. (2004). Effects of subject matter expert viewpoint on job analysis results. *Public Personal Management*, 33(1), 33-46.
- Von Heimburg, E. D., Rasmussen, A. K., & Medbo, J. I. (2006). Physiological responses of firefighters and performance predictors during a simulated rescue of hospital patients. *Ergonomics*, 49(2), 111-126.
- Von Heimburg, E., Medbo, J. I., Sandsund, M., & Reinertsen, R. E. (2013). Performance on a Work-Simulating Firefighter Test Versus Approved Laboratory Tests for Firefighters and Applicants. *International Journal of Occupational Safety and Ergonomics*, 19(2), 227-243.

- Wilkinson, David M., Rayson, Mark P., & Bilzon, James L. J. (2008). A physical demands analysis of the 24-week British Army Parachute Regiment recruit training syllabus. *Ergonomics*, 51(5), 649-662.
- Williams, A. N., Wilkinson, D.M., Richmond, V.L., & Rayson, M.P. (2012). Normal Pension Age for Firefighters. A Review for the Firefighters' Pension Committee.
- Williams-Bell, F. M., Villar, R., Sharratt, M. T., & Hughson, R. L. (2009). Physiological demands of the firefighter Candidate Physical Ability Test. *Med Sci Sports Exerc*, 41(3), 653-662.
- Willich, S. N., Lewis, M., Lowel, H., Arntz, H. R., Schubert, F., & Schroder, R. (1993). Physical exertion as a trigger of acute myocardial infarction. Triggers and Mechanisms of Myocardial Infarction Study Group. *N Engl J Med*, 329(23), 1684-1690.
- Williford, H. N., Duey, W. J., Olson, M. S., Howard, R., & Wang, N. (1999). Relationship between fire fighting suppression tasks and physical fitness. *Ergonomics*, 42(9), 1179-1186.
- Yang, J., Teehan, D., Farioli, A., Baur, D. M., Smith, D., & Kales, S. N. (2013). Sudden cardiac death among firefighters ≤ 45 years of age in the United States. *Am J Cardiol*, 112(12), 1962-1967.
- Zumbo, B. D. (2016). Standard-setting methodology: Establishing performance standards and setting cut-scores to assist score interpretation. *Appl Physiol Nutr Metab*, 41(6 Suppl 2), S74-82.